Molecular Cloning and Functional Characterization of a New Sucrose Transporter in Hexaploid Wheat (*Triticum aestivum* L.)

by

Kirandeep Kaur Deol

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

Department of Plant Science

University of Manitoba

Winnipeg

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

COPYRIGHT PERMISSION

Molecular Cloning and Functional Characterization of a New Sucrose Transporter in Hexaploid Wheat (*Triticum aestivum* L.)

By

KIRANDEEP KAUR DEOL

A Thesis/practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

Of

Master of Science

..

ABSTRACT

Deol, Kirandeep Kaur. M.Sc. The University of Manitoba, April, 2012. Molecular Cloning and Functional Characterization of a New Sucrose Transporter in Hexaploid Wheat (*Triticum aestivum* L.). M.Sc. supervisor: Dr. Belay T Ayele.

Current bioethanol production is predominantly based on starch from cereal crops such as corn and wheat which leads to increased competition for such crops between the food, feed and bioethanol industries. Thus, enhancing the sustainability of the bioethanol industry, at least in the short term, requires increased starch yield per seed or per unit area. Starch content of wheat grains is partly determined by sucrose transport efficiency from the source (the leaf) to the sink organ (the grain) which is partly mediated by sucrose transporter (SUT) proteins. Only one functional *SUT* (*TaSUT1*) gene of wheat is known to date. Using molecular gene cloning approaches, this thesis project identified the three homeologues of a new *SUT* (*TaSUT2*) gene from hexaploid wheat and demonstrated their functionality as sucrose transporters through heterologous expression in mutant yeast. Characterization of the spatio-temporal expression of *TaSUT2* suggested its role in phloem loading, unloading and sucrose retrieval.

ACKNOWLEDGEMENT

It is with gratitude that I acknowledge my heartiest thanks to Dr. Belay T. Ayele for his continuous support for my M.Sc. project and for providing me with wealth of information, motivation and guidance throughout my study period. I would also like to thank Drs. Claudio Stasolla, Anita Brûlé-Babel and Dana Schroeder for being on my advisory committee and providing me with their valuable suggestions and comments.

I would like to thank Ms. Chelsey McDougall for her kind help and sharing her technical experience. It is also a pleasure to thank all members of the Crop Technology Center in the Department of Plant Science for their support, valuable discussions and for the wonderful time we had together. I extend my appreciation to Husky Energy Inc. and Natural Science and Engineering Research Council of Canada (NSERC) for their financial support.

In addition, I express my gratitude to my family and friends. Specifically, I would like to thank my husband Manjitinder Deol for providing me support and motivation. I also extend my gratitude towards my loving son, Karman Deol, who is always my source of inspiration. Last but not least, I thank my parents for their never ending support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ADP	adenosine diphosphate
AGPase	adenosine diphosphate glucose pyrophosphorylase
ATP	adenosine triphosphate
BLAST	basic local alignment search tool
BTU	British thermal unit
cDNA	complimentary deoxyribonucleic acid
CPS	Canada Prairie Spring
DAA	days after anthesis
DAH	days after heading
DDGS	dried distillers grains with solubles
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
EDTA	ethylene diamine tetraacetic acid
HALC	high affinity low capacity
IPTG	isopropyl β-D-1- thiogalactopyranoside
K_m	michaelis-menten constant
LAHC	low affinity high capacity
LB	luria-bertani
mRNA	messenger ribonucleic acid
NADPH	nicotinamide adenine dinucleotide phosphate
NaH_2PO_4	sodium phosphate
NCBI	national centre for biotechnology information

ORF	open reading frame
PCR	polymerase chain reaction
RT-PCR	reverse transcriptase- polymerase chain reaction
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SS	synthetic complete media with sucrose as the sole carbon source
SUC	sucrose carriers
SUT	sucrose transporter
SWS	Soft White Spring
U	units
UDP	uridine diphosphate
YPD	yeast peptone dextrose

FORWARD

This thesis is written in manuscript style. A general introduction about the research project and a literature review precedes the manuscript. An abstract, introduction, materials and methods, results and discussion form a complete manuscript. The manuscript is followed by a general discussion and conclusions, a list of references and appendices.

1.0 INTRODUCTION

The finite availability of fossil fuels, increased concerns about global warming associated with their continued use and the ever fluctuating price of oil in the world market have triggered the need to develop renewable sources of energy such as biofuels. Currently, starch and sugar crops serve as the main feedstock for the production of bioethanol, a form of biofuel widely used as liquid fuel for transportation purposes across the world (Ajanovic, 2010), though efforts are underway to develop efficient and cost effective methods of producing bioethanol from cellulosic biomass. Wheat grains are the most important raw material for starch based bioethanol production in Western Canada (Olai et al., 2004). However, the rising demand for wheat as a resource for bioethanol production is leading to increased competition for this crop between the food, feed and bioethanol industries. Therefore, in order to meet the increasing demand for wheat and to make the bioethanol industry sustainable at least in the short term until bioethanol production from cellulosic biomass become economically feasible, there is a need to increase grain starch yield per plant or per unit area. This has triggered a wide range of research efforts to investigate the factors determining starch yield in wheat.

Starch yield in the grains of cereal crops is determined mainly by the efficiency of photosynthesis and translocation of photosynthates from the source (mature leaves) to the sink (grain) tissues. Since sucrose is the major form of photoassimilates that can be translocated from source to sink organs, starch accumulation in the endosperms of cereal grains can be enhanced by increasing the sucrose concentration made available to the developing grains, which further depends on the efficiency of sucrose transport from the site of its synthesis to the endosperm. Sucrose transporter (SUT) proteins play a major

role in this process (Truernit, 2001). Genes encoding SUT proteins have been isolated from a number of cereal crops that are widely cultivated around the world; including five SUTs from rice, five from maize, two from barley and one from wheat. The SUTs of cereal crops are reported to be mainly associated with grain filling as the filial generation has no symplastic connections with the maternal tissues in the seeds (Weschke et al., 2000). Previous reports have reported the enhancement of crop yield through manipulation of the activity of SUTs. For example, ectopic expression of *SUT*s has been demonstrated to increase yield in crop plants such as potato (Leggewie et al., 2003) and pea (Rosche et al., 2002). These results indicate the potential of using *SUT* genes as a tool to increase sucrose transport from the source (leaf) to the sink (grain) tissues of cereal crops, and thereby enhance grain starch yield. To date, only one *SUT* gene (*TaSUT1*) has been identified from wheat (Aoki et al., 2002). This thesis reports the cloning and characterization of three homeologues of a new sucrose transporter gene, designated as *TaSUT2A, TaSUT2B* and *TaSUT2D*, from hexaploid wheat (*Triticum aestivum* L.) cv. AC Andrew. The findings of this research have the potential to further our understanding of carbon partitioning and provide a tool for the development of high yield and high starch content wheat that is in demand by the bioethanol industry in Western Canada.

2.0 LITERATURE REVIEW

2.1 Background

Ever since the industrial revolution, human beings have been dependent on fossil fuels as the source of energy. The current global energy demand is estimated to be approximately 400 quadrillion British Thermal Unit (BTU) per annum, and is predicted to rise to 770 quadrillion BTU by the year 2035, according to the 2011 US Department of Energy report. Fossil fuels account for 350 quadrillions BTU or 88% of the world's energy consumption. Given that fossil fuels are non-renewable and limited in supply, there is a threat that they will not last longer than 100 to 200 years (U.S. Energy Information Administration, 2011). Furthermore, the greenhouse gases released from fossil fuel use are causing global warming as they trap 90% of the radiation emitted by earth's surface (Srinivasan, 2008).

Biofuels serve as an alternative to fossil fuels as they are renewable and release fewer greenhouse gases (Hahn-Hagerdal et al., 2006). At present biofuels are produced mainly from starch and sugar crops such as corn, wheat and sugarcane (Ajanovic, 2010). Starch contains a large number of glucose molecules joined together by glycosidic bonds, and acts as a major storage reserve in the grains of cereal crops. Its biodegradable and renewable properties make this form of carbohydrate an excellent resource for bioethanol production. The deposition of starch in cereal grains can be enhanced by using a number of strategies such as by manipulating starch biosynthetic pathway and increasing the amount of sucrose made available from source tissues, which in turn is dependent on the activity of proteins that mediate its active transport into the developing grains.

2.2 Biofuels

2.2.1 Introduction to biofuels

Biomass derived fuels are known as biofuels and can be found in different phase/states, such as solid biomass, liquid fuels or biogases. They are classified into four types based on their generation. First generation biofuels are derived from sugar, starch, vegetable oil and animal fats, and comprise bioethanol, biopropanol, biobutanol, biodiesel and biogas. Edible parts of food crops mainly the grains of cereal crops form the main feedstock for the production of first generation biofuels (Demirbas, 2009; Rajendrani, 2011). The second generation biofuels, which consist of cellulosic bioethanol are produced from the non-edible part of crop plants such as stalks of wheat and corn, and biomass derived from non-food/energy crops such as Miscanthus and wood (Simpson-Holley and Evans, 2007; Naik et al., 2010). Therefore, they have an advantage over the first generation biofuels as they help save food crops for human and livestock consumption. The third generation biofuels are produced from algae, and thus sometimes referred as algal fuel. Algae are considered as suitable source of biofuel since they produce high yield per unit input and are harmless to the environment. The production of fourth generation biofuels mainly involves metabolic engineering of microorganisms (Lu et al., 2011).

2.2.2. Advantages of biofuels over fossil fuels

Generation of fossil fuels takes thousands of years (Demirbas, 2009). This, along with the unpredictability of fuel prices, has triggered the need to find a more sustainable source of energy. Green plants are considered good alternatives as they can be regenerated many times, and thus serve as a renewable source of energy (Demirbas, 2009). One of the

major challenges associated with the use of fossil fuels is the emission of greenhouse gases that cause global warming and environmental pollution. As biofuels emit less greenhouse gases, their use has an advantage over that of fossil fuels with this respect. There is virtually no net release of carbon dioxide associated with the use of biofuels as carbon released from their burning is utilized by plants to produce carbohydrate and support their growth (Hill et al., 2006; Demirbas, 2007). Biofuels are also less toxic than fossil fuels since they are produced from biodegradable matter, causing no environmental pollution or health hazard. Furthermore, the production of biofuels can help to generate additional farm income and tackle waste storage problem as they can also be produced from plant wastes such as corn stalks. Finally, the generation and use of fuels derived from plant biomass can lead to a decrease in dependence on foreign fossil fuel sources and provide better energy security for most of the nations in the world as plants can easily be produced by almost all countries.

2.2.3 Bioethanol

Bioethanol, which is widely used as liquid biofuel for transportation purposes (Balat and Balat, 2009), is produced primarily from starch and sugar deposited in the storage organs of crops such as wheat, corn, potato and sugarcane, and to some extent from cellulosic biomass of crops, grasses and trees. At present bioethanol is used mainly as a gasoline additive to increase the gasoline octane rating, thereby reducing the amount of carbon monoxide produced by the vehicle. Bioethanol is an oxygenated fuel that comprises 35% oxygen, making it one of the best alternatives to fossil fuels as this property of biofuels is

important to decrease the emissions of particulates and NO_x (Demirbas, 2005). In addition, it can be used in its pure form as a substitute of gasoline in petrol engines.

The USA and Brazil are among the leading producers of bioethanol, and in 2009 they accounted for 86% of world's total bioethanol production, which was estimated to be 73.9 billion litres (L). Global bioethanol production reached 85.8 billion L in 2010, and was estimated to increase to 88.7 billion L by 2011, out of which the USA alone was expected to produce 51 billion L (Licht, 2011). About 95% of USA's bioethanol is produced from corn, whereas sugarcane is the main bioethanol feedstock in Brazil. In Canada, bioethanol is mainly produced from starchy crops like corn and wheat (Balat and Balat, 2009) and its production is expected to reach 1350 million L by 2011 and 1375 million L by 2012 (Dessureault, 2011). Wheat grains are the major feedstock in Western Canada (Saunders et al., 2011).

2.2.4 Comparison between starch and cellulose based bioethanol

Bioethanol production is mainly based on three types of raw materials - simple sugars, starch and cellulose (Demirbas, 2005; Naik et al., 2010). Its production involves the breaking down of complex carbohydrates into simple sugars, which are further fermented into ethanol. Starch based bioethanol is mainly produced from corn kernels and wheat grains, thus it is sometimes referred as grain alcohol. The grains or kernels are first ground by milling, then common enzymes such as amylases are used to break the starch into glucose, which is fermented by yeast to form ethanol. In contrast, the production of bioethanol from cellulose is much more complicated and slower than that of starch (Gray et al., 2006; Crowe, 2008). Plant cell walls, comprised of polysaccharides such as

cellulose and hemicelluloses, as well as lignin mainly form the plant biomass from which bioethanol is produced. Similar to starch, cellulose is also composed of long chains of glucose; however, it has a different structural configuration. In starch, glucose molecules are linked by α-1, 4 glycosidic linkage (amylose) and by both α-1, 4 and α-1, 6 glycosidic linkages (amylopectin), whereas in cellulose they are linked by β-1, 4 glycosidic bonds which makes cellulose more crystalline and compact (Demirbas, 2005; Gray et al., 2006; Naik et al., 2010).

Cellulose and hemicellulose are surrounded by a lignin layer, which is a complex polymer structure made from phenolic groups. Lignin has a negative impact on cellulosic biofuel production as it adsorbs the hydrolytic enzymes used to degrade the polysacchrides, thus interfering with their access to the cellulose from which they release monosaccharides. Since lignin is hard to break down due to the strong bonds that hold its polymers, heat or strong chemicals such as acids and ammonia are required to degrade the lignocellulosic material and free up the cellulose and hemicelluloses (Sanderson, 2011). Furthermore, lignin degradation products are known to inhibit the fermentation step in the process of biofuel production (Li et al., 2008; Fu et al., 2011). All these factors contribute to the high cost of bioethanol production from lignocellulosic feedstock.

Unweaving the cellulose component and breaking down of the hemicelluloses into glucose and pentose sugars including xylose, mannose, arabinose and galactose requires pretreatment with acids or enzymes. The pentose sugars, however, make the fermentation process more complicated as compared to that of starch based feedstock (Demirbas, 2005). To overcome this problem and facilitate the fermentation process, specialized and engineered microbes are being used. However, incomplete conversion of

the biomass to ethanol along with the slow and expensive enzymatic reactions still make bioethanol production from cellulosic feedstock more costly than that derived from starch. Efforts are underway to develop methods to release the sugars deposited in plant cell wall polysaccharides and produce bioethanol in an efficient and cost effective way (Carroll and Somerville, 2009; Naik et al., 2010).

2.2.5 Concerns associated with bioethanol production

There are some concerns associated with the expansion of the biofuel industry, mainly with that of first generation biofuels that compete for starch with other uses such as food and feed. Given that it is renewable, the increasing demand for bioethanol will lead to the need for large amounts of arable land, leaving less land available for growing crops for food and feed purposes, which will ultimately lead to a rise in food prices. Therefore, as starch and sugar crops will remain as the main source of bioethanol at least in the short term, increasing their productivity per plant or unit area is considered as the best alternative to make more land available for food and feed production (Smith, 2008).

2.2.6 Bioethanol production in Canada

The cereal grain based bioethanol industry is expanding in Canada. Corn is the major bioethanol feed stock in the eastern part of the country (Olai et al., 2004). The western provinces lack the environmental conditions such as the growing degree days that are needed for growing corn (Olai et al., 2004; Austin, 2009), as a result wheat, barley, oat and triticale are widely grown in this region and serve as feedstock for bioethanol production. Wheat is the major cereal crop in Western Canada and the starch content of

its grain is the highest among cereal crops (starch accounts for about 61% of wheat grain dry weight as compared to 42% and 31% of barley and oat, respectively), thus wheat is used as the main resource for bioethanol production (Strategic Vision Consulting Ltd [SVC], 2007). The suitability and choice of a wheat variety for bioethanol production depends on its starch yield. Thus, availability of high starch content wheat cultivars make starch based bioethanol industry more sustainable.

2.2.7 Bioethanol plants in Western Canada

A total of 15 operational bioethanol plants are present in Canada, and eight of these plants are located in the western part of the country. In Alberta, bioethanol plants mainly rely on wheat as a feedstock for ethanol production, though corn, barley, rye and triticale are used to a certain extent. In contrast, bioethanol plants in Manitoba and Saskatchewan mainly use grains of wheat and corn as feedstock. Among the bioethanol plants found in Western Canada, the Minnedosa bioethanol plant in Manitoba and the Lloydminster bioethanol plant in Saskatchewan, which are owned by Husky Energy Inc., have the highest capacity of bioethanol generation, accounting for about 130 million L per year (Canadian Renewable Fuel Association, 2010).

2.3 Wheat production

2.3.1 Origin and history of hexaploid wheat

Wheat cultivation started about 10,000 years ago, when humans started practicing agriculture, using diploid (einkorn) and tetraploid (emmer) forms of wheat that are considered to be the primitive relatives of today's cultivated wheat crops. Wheat

cultivation then spread to the Near East, where hexaploid wheat was firstly discovered about 9000 years ago (Shewry, 2009). Wheat belongs to the genus *Triticum* and the family Gramineae. On the basis of their ploidy level, *Triticum* species are grouped into three clusters: the diploid species that have 14 chromosomes and the tetraploid and hexaploid species that contain 28 and 42 chromosomes, respectively. The crosses between diploids followed by chromosome doubling, gave rise to tetraploid wheat. Tetraploid wheat further hybridized with a diploid species result in the hexaploid wheat (Matsuoka, 2011). Wheat has three different sets of genomes, named as A, B and D. The diploid wheat species has an AA genome, whereas the tetraploid and hexaploid species contain AABB and AABBDD genomes, respectively (Feldman, 1997; Curtis, 2002; Shewry, 2009). The diploid species, *Triticum monococcum* (Einkorn wheat) with the AA genome, and the tetraploid wheat, *Triticum dicoccum* (Emmer) with the AABB genome, are among the earliest cultivated wheat species (Curtis, 2002; Shewry, 2009). *Triticum monococcum* and *Triticum urartu* are considered as valid biological species and thought to be the donors of the A genome to both tetraploid and hexaploid wheat, whereas *Triticum tauschii* also known as *Aegilops tauschii*, is considered as the D genome donor to hexaploid wheat (Curtis, 2002). *Aegilops speltoides* is regarded as the B genome donor of both tetraploid and hexaploid wheat (Gill and Friebe, 2002; Shewry, 2009).

2.3.2 Wheat production in Western Canada

The cultivation of wheat in Manitoba is associated with the Selkirk settlers, who came from Scotland in 1812 and colonized the territory granted to them by the Hudson's Bay Company. They established the Red River settlement at the junction of Red and

Assiniboine rivers and grew the first wheat, a winter wheat they brought from their native land. A year later the first spring wheat was grown in Western Canada. In 1857, the Mennonites from Russia arrived in Manitoba along with many wheat varieties produced in their country. This has led to the expansion of wheat cultivation in Western Canada. For example, wheat production rose from 1.5 million metric ton (MT) in 1904 to 5.6 million MT in 1913 (Buller, 1919).

Canada is a major producer and exporter of wheat, ranking sixth in the world (http://www.gov.mb.ca/trade/globaltrade/agrifood/commodity/). The Prairie Provinces of Western Canada including Manitoba, Saskatchewan and Alberta, alone account for 95 percent of total wheat produced in Canada (Curtis, 2002). Canada Western Red Spring (CWRS) is the main wheat class grown in western Canada and accounts for 80% of wheat production, while Canada Western Amber Durum (CWAD) is the second most important wheat grown in the region. The current total wheat production in the Prairie provinces is estimated at 21.6 million MT in 2011, of which Spring wheat accounts for about 17.1 million MT while durum wheat about 3.7 million MT (Canada Wheat Board, 2011).

2.3.3 Wheat as a feedstock for bioethanol in Western Canada

Canadian Prairie Spring (CPS) red and white, Canada Western Soft White Spring (CSWS), Canada Western Red Winter (CWRS) and Canada Western General Purpose (CWGP) wheat cultivars are the main choices for bioethanol production. Soft white and CPS wheat are preferred as they produce less dried distillers grains (DDGS) than the hard red spring varieties that produce more DDGS than bioethanol (SVC, 2007). Among the

currently registered wheat varieties in Western Canada, the soft white spring wheat cultivar (cv.) AC Andrew stands first in terms of yield and starch content (Austin, 2009). It has 15% to 19% higher yield than the low-protein Soft White Spring (SWS) wheat. The other important feature that makes cv. AC Andrew suitable for bioethanol production is its low protein content per grain because low protein and high starch yielding wheat cultivars are able to produce more grain, and thereby more bioethanol per unit area as compared to cultivars with high protein and low starch yield (Rharrabti et al., 2001; Secan, 2009).

2.4 Starch formation in plants

The current research with respect to bioethanol crops is mainly focused on enhancing their productivity to maintain the sustainability of bioethanol industry. Photosynthesis is the process by which green plants convert light energy into carbohydrates from CO_2 and water. It takes place in the mesophyll cells of leaves, where chloroplasts, which contain the light harvesting pigments such as chlorophyll, are located (Taiz and Zeiger, 2010a). The production of photoassimilates in the source tissues and their partitioning, translocation and ultimate accumulation in the sink organs are of immense importance in crop production. Starch, which is currently used as a substrate for bioethanol production, is the major storage reserve deposited in the grains of cereal crops. Starch accumulation in these storage organs is strongly dependent on source-sink relationships, which are defined mainly by the pathway through which photosynthate translocation occurs (Emes et al., 2003). Thus, one approach to increasing starch yield is through manipulation of the

translocation of photosynthates from the source to the sink organs (Martin and Smith, 1995; Smith, 2008).

2.4.1 Photosynthesis

The carbon compounds formed in the photosynthetic reaction are used to meet the energy requirements of all cellular processes in plants. Photosynthesis involves two reactions: the light reaction and carbon fixation. The light reaction involves the formation of the energy carrier molecules, adenosine triphosphate (ATP) and nicotinamide adenine dinucleotide phosphate (NADPH), whereas carbon fixation (also called the Calvin cycle) utilizes ATP and NADPH to fix CO_2 and produce triose phosphate, which is used as a precursor for starch and sucrose synthesis (Taiz and Zeiger, 2010b). During the active photosynthetic phase, the triose phosphate is used for the formation of chloroplast-stored starch, which is referred as transient starch. In the dark, when there is no photosynthesis, the transient starch is converted into sucrose so as to ensure the continuous availability of photosynthates to the whole plant. The sucrose produced during the active photosynthetic or dark phase is either stored in the vacuole or translocated into the sink organ where it is used as a substrate for the synthesis of storage compounds such as starch that is deposited in the endosperm of developing wheat grains (Trethewey and Smith 2000).

2.4.2 Transitory starch and sucrose synthesis compete with each other

Transitory starch synthesis in the chloroplast and sucrose synthesis in the cytosol are two competing reactions. The rate and type of reaction taking place depends on the cellular concentrations of triose phosphate and orthophosphate. The phosphate translocator acts

as an antiporter and controls the movement of triose phosphate and orthophosphate between the cytosol and the chloroplast. Each triose phosphate exported from chloroplast to cytosol is accompanied by intake of orthophosphate into the chloroplast, thus maintaining their pool within the stroma. The exported triosphosphate is used for sucrose synthesis in the cytosol, resulting in the release of orthophosphate which moves into the chloroplast, allowing for continuous flow of triosphosphate into the cytosol. Non availability of inorganic phosphate due to decreased sucrose synthesis at the times of low sink demand limits the export of triosphosphate into the cytosol, making the triosephosphate more available for transitory starch synthesis in the choloroplast (Heldt et al., 1977; Tretheway and Smith, 2000). Transitory starch is mainly synthesized during the daytime, acting as temporary storage for starch that will be degraded during the night for translocation into sink organs. The accumulation of transitory starch in the chloroplast is low during the early hours of the day when there is a high sink demand but its level eventually increases as the sucrose level in the leaves rises due to low sink demand. The extent of transitory starch accumulation also varies with species. Many species in the gramineae family, such as wheat do not store starch in the chloroplast, rather they store fructans in the vacuoles of their mesophyll cells (Tretheway and Smith, 2000).

2.4.3 Starch as a storage reserve

Sucrose, as the major product of photosynthesis and translocated form of carbohydrate, can be stored in the vacuole or be transported via phloem to storage organs where it is used for starch synthesis (Wolosiuk and Pontis, 1974; Keeling et al., 1988). The rate of starch synthesis in the storage organs mainly depends on the rate at which sucrose enters

into these tissues. The very first step of starch synthesis involves the formation of UDP glucose and fructose from sucrose. This process is mainly catalyzed by two enzymes, sucrose synthase and invertase, in the cytosol. The resulting UDP glucose and fructose are then converted into hexophosphates (glucose-1-phospahate and glucose-6-phosphate), which are imported into the amyloplast. Starch synthesis in the amyloplasts proceeds with the help of the plastidial isoform of phosphoglucomutase, which catalyzes the conversion of glucose-6-phosphate into glucose-1-phosphate. ADP-glucose is then synthesized from glucose-1-phosphate by the action of ADP-glucose pyrophosphorylase (AGPase). Starch synthases and branching enzymes further catalyse the polymerization reactions that lead to the formation of starch (Emes et al., 2003). Starch synthesis in cereal endosperms has a unique feature of having AGPase activity in both plastids and cytosol, whereas in non-cereal species and other cereal organs, AGPase activity is always plastidial. Thus, in cereal grains such as wheat, ADP glucose is mostly synthesized by cytosolic AGPase in the cytosol after which it is transported with the help of specific sugar nucleotide transporters into the amyloplast, where it is used as a substrate for starch synthesis (Smith, 2008; Tomlinson and Denyer, 2003).

2.5 Source-sink relationships

Source tissues comprise any organ that has the ability to produce enough photosynthates for export to sink organs, which comprise of any non-photosynthetic organ or any organ that has less capacity to produce the photosynthates required for its growth and development or storage (Williams et al., 2000; Emes et al., 2003). The transport of photoassimilates from source to sink organs takes place in the phloem and follows a

specific developmental and anatomic pattern. All source tissues cannot supply photosynthates to all sink organs of a given plant. The partitioning of photosynthates from a certain source organ into a specific sink tissue is governed by many factors, including the ability of the sink organ to unload nutrients from the importing sieve elements, sink strength, proximity of the sink organ to the source tissue, availability of direct vascular connections between the source and sink tissue, and developmental stage of the sink organ (Simmons, 1987; Patrick and Offler, 2001).

The proximity or closeness of sink tissue to source tissue is the major determining factor for the partitioning of photosynthates. Leaves in the upper layer of a crop canopy usually supply photosynthates to growing shoot tips or young leaves, while those in the lower layer supply the root. The leaves in the middle layer of the crop canopy usually translocate photosynthates in both the directions (Taiz and Zeiger, 2010c). In wheat, the flag leaf, which is closer to the spike than any other leaf, is the major contributor of photoassimilates to developing grains (Araus and Tapia, 1987). The developmental stage of a specific sink organ also determines its strength to accumulate photosynthates. For example the root and shoot are the major sink organs during the vegetative growth of a plant whereas fruits and developing grains become the dominant sink organs during the reproductive phase of plant development. Moreover, the presence of direct vascular connections with the source tissue contributes to the strength of a sink organ.

2.5.1 Photosynthate transport through the phloem

Sucrose is the major form of carbohydrate translocated through phloem to various sink organs (Lim et al., 2006). This is mainly because of its non-reducing nature, which makes

it chemically stable as compared to reducing sugars. Moreover, the high osmotic pressure created by sucrose per carbon atom in the phloem sap is also responsible for its high translocation efficiency within sieve tube elements (Kuhn et al., 1999; Lemoine, 2000). Phloem is the main tissue responsible for the translocation of photosynthates in the form of sucrose from mature source tissues to developing immature source and storage sink organs. Phloem tissues consist of sieve elements (SE) and companion cells (CC) as their main components. The SEs are specific cells in the phloem that mediate long distance transport of sucrose. They are joined together to form sieve tubes. Sieve plates that have large plasma membrane lined pores are formed on the end walls of the sieve tube elements. These pores act as open channels for photoassimilate transport between cells. The sieve tube elements are connected with the companion cells by numerous intercellular connections called plasmodesmata, which are plasma membrane lined narrow channels that act as a pathway for transporting photosynthates. Normally each sieve tube element is associated with one or more companion cells that have high metabolic activity and act as a source of proteins and energy for the sieve elements, which lose most of their organelles including the tonoplast, nuclei, ribosomes and microtubules during their development (Kuhn et al., 1999; Lemoine, 2000; Williams et al., 2000).

2.5.2 Mechanism of phloem loading and unloading

Sucrose is delivered into the sieve element-companion cell (SE-CC) complex mainly by two pathways: symplastic and apoplastic. After its synthesis in the mesophyll cells of the leaves, sucrose may enter the SE-CC complex through the symplastic pathway, which

involves the use of plasmodesmatal connections. Alternatively, it may follow the apoplastic pathway in which sucrose released from mesophyll cells is actively loaded into the SE-CC complex with the help of plasma membrane localized sucrose H^+ symporters, also known as sucrose transporter (SUT) proteins (Kuhn et al., 1999; Truernit, 2001; Sauer, 2007). Therefore, the SUTs are needed for active transport of sucrose across the plasmamembrane into the apoplast and also for transport of sucrose out of vacuoles across the tonoplast (Ayre, 2011). Apoplastic phloem loading requires the use of energy as the movement occurs against the concentration gradient because sucrose is present in higher concentration in the phloem than in the apoplast (Braun and Slewinski, 2009). This energy is generated by proton pumping ATPases that are co-localized with the SUTs in the plasma membrane of sieve elements and companion cell (Harrington et al., 1997). The ATPases are able to establish a high proton gradient in the apoplast by pumping protons out of the cell and the resulting proton gradient energy is used for sucrose transport into the SE-CC complex via the SUTs (Harrington et al., 1997).

The number of plasmodesmatal connections between the minor vein SE-CC complex and its surrounding cells varies with species, and is the main factor that determines the type of sucrose transport pathway operating. For example, the SE-CC complex is almost symplastically isolated in many crop species such as tobacco and potato; therefore the apoplastic pathway is the main route of sucrose transport for these crops. However, in species such as Arabidopsis, even if there are numerous plasmodesmatal connections between the SE-CC complex and its surrounding cells, sucrose is actively loaded into the phloem from the apoplast with the help of SUTs (Truernit, 2001). Both the symplastic and apoplastic pathways can operate in the same

species depending on tissue type and developmental stage (Williams et al., 2000; Truernit, 2001).

Phloem unloading of sucrose in the sink organs also takes place via the apoplastic and symplastic pathways. Following its long distance transport in the phloem, sucrose exits the sieve elements of sink tissues, which is called sieve element unloading, and enters the sink cells by the mechanism known as post-sieve element transport. The sugar imported into the sink cells is used for storage or metabolic activities. In cereal grains, the maternal and filial generations are isolated apoplastically, thus the post sieve element transport of sucrose from the apoplast across the aleurone layer into the endosperm, where it is used for starch synthesis, is mediated by SUTs. Alternatively, the sucrose is first hydrolyzed by cell wall invertase into glucose and fructose, which are subsequently taken up into the sink cells by monosaccharide transporters (Oparka, 1990).

2.6 Sucrose transporters

All sucrose transporter (SUT) proteins belong to the glycoside-pentoside-hexuronide (GPH) cation symporter family, which is a member of the major facilitator superfamily (Chang et al., 2004; Sauer, 2007). The major facilitator superfamily, is one of the largest families of membrane transporters, may work as uniporters, symporters or antiporters (Pao et al., 1998). Members of the GPH gene family encode highly hydrophobic proteins with 12 transmembrane-spanning helices (Truernit, 2001). The N- and C-termini of these proteins are cytoplasmic. The alignment of the first and second halves of the SUTs suggests that these proteins may have evolved from one or more ancestral transporters via gene duplication and fusion (Sauer, 2007). A diethyl pyrocarbonate-sensitive histidine

residue located at position 65 of the SUT protein from Arabidopsis is conserved across all other SUTs identified so far. The His-65 residue is thought to be involved in sucrose binding during transportation (Scofield et al., 2002).

2.6.1 Classification of SUTs

SUTs have been isolated from both monocot and dicot species. The first SUT protein was isolated from spinach, SoSUT1 (Riesmeier et al., 1992). To date, nine SUTs have been identified in Arabidopsis (Eckardt, 2003), five in rice (Aoki et al., 2003), two in barley (Weschke et al., 2000), one in wheat (Aoki et al., 2002) and many more from other dicot and monocot species (Eckardt, 2003). On the basis of their kinetic properties, the SUT proteins are categorized into two transport systems: high affinity low capacity (HALC) and low affinity high capacity (LAHC). The HALC SUTs have higher affinity for sucrose as they can uptake sucrose present at low concentration. The HALC transport system plays a role in phloem loading and retrieval of the sucrose that diffuses out of vascular tissues during transport (Aldape et al., 2003). This transport system is necessary to maintain a continuous flow of sucrose from source tissue and maintains a solute concentration gradient in the phloem. Members of the *SUT1* gene family from dicots generally encode high affinity SUTs that have Km values of 0.3 to 2.0 mM, and fall under the HALC category. For example, SUT1 of potato is a high affinity SUT with a Km value of 1 mM (Riesmeier et al., 1993). The LAHC SUTs have higher Km values and lower affinity for sucrose, as they uptake sucrose only when it is present at high concentration. This transport system is mainly important for phloem loading (Weise et al., 2000; Aldape et al., 2003). The cytoplasmic N- terminus of SUT proteins may be

involved in determining their affinity for sucrose substrate (Schulze et al., 2000). However, since it is cytosolic, the N-terminus cannot bind sucrose directly but may influence affinity for the sucrose through intramolecular interactions with other cytoplasmic domains (Schulze et al., 2000).

All the known SUTs have been categorized into three major groups by Aoki et al (2003) based on similarity of their amino acid sequences: type I (SUT1), type II (SUT2/3), and type III (SUT4). The Type I SUTs consist of mainly dicot sucrose transporters, which are high affinity/low capacity SUTs with Km values ranging from 139 μM to 1.5 mM. Many high affinity SUTs from dicot species such as SUT1 proteins from spinach (SoSUT1), Arabidopsis (AtSUT1), potato (StSUT1) and tomato (LeSUT1) fall under this category. Members of Type III SUTs exhibit low affinity/high capacity nature with Km values ranging from 5 to 6 mM. Some SUTs from monocots such as SUT2s from barley (HvSUT2) and rice (OsSUT2) as well as low affinity dicot SUTs such as SUT4s from potato (StSUT4) and tomato (LeSUT4) belong to this group. The type II SUTs are divided into two different sub groups, leading to classification of the plant SUTs into 4 groups (Sauer, 2007). One group includes SUTs from monocots like SUT1s from wheat (TaSUT1), rice (OsSUT1) and maize (ZmSUT1), and the second group is composed of very low affinity SUTs from dicots such as SUT2s from tomato (LeSUT2) and Arabidopsis (AtSUT2). These SUTs lack transport activity and structurally resemble yeast sugar sensors, SUCROSE NONFERMENTING 3 (SNF3) and RESTORES GLUCOSE TRANSPORT 2 (RGT2; Barker at al., 2000), thus are believed to act as sugar sensors. The SNF3 and RGT 2 sensor proteins resemble glucose transporters but lack transporting activity. The SNF3 is a yeast glucose sensor protein that

senses low glucose levels while high glucose levels are sensed by RGT2, making the yeast able to react according to a changing environment and regulate the expression of genes encoding other glucose transporters (Barker at al., 2000).

Owing to the completion of the draft genome sequences of maize, sorghum and Brachypodium, SUTs have increased in number and recently been classified into 5 groups (Braun and Slewinski, 2009). The monocot SUTs previously categorized as Type II SUTs (Aoki et al., 2003) have been split into two different groups, namely group 1 (consists of SUT1 proteins from monocot species such as TaSUT1, ZmSUT1, OsSUT1, OsSUT3), and group 5, which mainly consists of the SUT5 proteins from monocots including OsSUT5, SoSUT5 and ZmSUT5. The high affinity SUT1 proteins from dicots, classified as Type I SUT proteins by Aoki et al. (2003), form group 2, according to the new classification. Both monocot and dicots SUT proteins are present in group 3, and the proteins in this group were classified as Type II SUTs by Aoki et al. (2003). The type III SUTs are reclassified as group 4 SUTs (Braun and Slewinski, 2009; Slewinski et al., 2009).

2.6.2 Nomenclature of SUTs

The SUT proteins are named according to the order in which they were isolated (Braun and Slewinski, 2009). For example, the first one to be isolated from Arabidopsis was named as AtSUT1. In order to minimize confusion with naming and determine their phylogenic relationship, the current nomenclature of SUTs is, however, mainly based on sequence similarity with the already identified SUTs. As rice SUTs were the first to be identified from monocots, newly identified monocot SUTs are named according to their

sequence similarity with the known rice SUTs. All known SUTs act as sucrose proton symporters and are energy dependent, and the proton pumping plasma membrane localized ATPases provide the energy required for transporting sucrose against its concentration gradient (Takeda et al., 2001; Braun and Slewinski, 2009).

2.6.3 Localization of SUTs

The SUT proteins are known to be localized on the sieve elements and companion cell plasma membranes. The transcripts of *SUT*s are however synthesized in the companion cells, as ribosomes and nuclei are absent from mature sieve elements (Kuhn et al., 1997). The dicot SUT1 proteins such as those from potato and tomato are localized in the plasma membrane of phloem sieve elements (Kuhn et al., 1997; Kuhn 2003; Barker et al., 2000) whereas sucrose carrier 2s (SUC2s) from Arabidopsis and Plantago are located in the plasma membrane of phloem companion cells (Stadler et al., 1995; Stadler and Sauer, 1996). Furthermore, monocot SUTs from Aoki's SUT2/3 (Type II) group such as SUT1s from barley and wheat have been localized in the plasma membrane of sieve elements (Weschke et al., 2000; Aoki et al., 2004), whereas rice SUTs (OsSUT1 and OsSUT3) are found to be present in the plasma membranes of both sieve elements and companion cells (Scofield et al., 2007; Kuhn and Grof, 2010). It has been shown recently that members of the SUT4 group such as SUT2 from barley and SUT4 from Arabidopsis (Endler et al., 2006) and Lotus (Reinders et al., 2008) are localized in the vacuolar tonoplast. Such vacuolar SUTs are proposed to be mainly responsible for sucrose efflux from the vacuole (Kuhn, 2011). However, the SUT4 from Arabidopsis is also found in chloroplast fractions (Rolland et al., 2003; Kuhn, 2011). In general, the SUTs from monocots have

lower affinity for sucrose as compared to those from dicots, and no monocot species has been found to encode SUTs like the dicot SUTs categorized in type 1 SUTs, which are regarded as SUTs with very high affinity for sucrose (Carpento et al., 2005).

2.6.4 Functions of SUT proteins

Phloem localized SUT proteins are mainly responsible for loading sucrose into the SE-CC complex and maintaining high sucrose concentration in the phloem to sustain its continuous supply to the sink organs (Sauer, 2007). Therefore, SUTs perform the special function of phloem loading when expressed in the source tissues. The importance of SUTs in phloem loading has been indicated by studies that involve *in situ* hybridization based gene expression analyses, which showed that *SUT*s are expressed only in the vascular tissues of source but not sink leaves (Riesmeier et al., 1993; Truernit and Sauer, 1995; Sauer, 2007). The expression of *SUT* in developing leaves begins at the tip and proceeds towards the base during its transition from sink to source organ (Sauer, 2007). The role of SUTs in phloem loading has also been demonstrated by antisense expressions of *SUT1* from potato (Riesmeier et al., 1994) and tobacco (Burkle et al., 1998), which resulted in accumulation of soluble carbohydrates and starch in the source leaves as export of sucrose was substantially reduced. Reduction in the export of sucrose was shown to affect whole plant growth and development, resulting chlorosis in leaves and inhibition of sink development (Sauer, 2007).

The phloem localized SUTs are also believed to perform the additional function of retrieving sucrose that leaked out of the sieve elements along the translocation pathway (Williams et al., 2000). For example, wheat SUT1, which is present in the

phloem of internodes, has been implicated in playing the role of retrieving sucrose leaked into the apoplast (Aoki et al., 2004). The SUTs that assist in the retrieval of sucrose also catalyze the export of sucrose from the source tissues and its uptake into the sink organs. It has been shown, for example, that the *SUT1* of potato is expressed in stems and sink tissues where it is believed to perform the role of sucrose retrieval and sucrose uptake. This gene is also found to be highly expressed in phloem of minor veins of the leaf, where it contributes mainly to phloem loading (Riesmeier et al., 1993; Truernit, 2001).

The SUT proteins are also detected in sink tissues, where they act as phloem-unloaders. During phloem unloading, sucrose is unloaded from the sieve elements of the sink organs, and then taken up by sink cells via post sieve element transport to be destined for storage or metabolism (Taiz and Zeiger, 2010c). For example, in cereal grains where maternal and filial generations are isolated from each other apoplastically (Weschke et al., 2000; Patrick and Offler, 2001), the SUTs play an important role in transporting sucrose to the endosperm, where it will be readily available for starch synthesis.

Some SUTs have been proposed to function as sucrose sensors. It has been hypothesized that *SUT2* of tomato (LeSUT2) can sense the concentration of sucrose, thereby regulating the expression of high and low affinity *SUT*s. This hypothesis is mainly based on the structural similarity of LeSUT2 with yeast sugar sensors such as SNF3 and RGT2, and its colocalization with high (LeSUT1) and low (LeSUT4) affinity SUTs in the sieve elements. Furthermore, similar to those found in the yeast sugar sensors, SNF3 and RGT2, the cytoplasmic domains of SUT2 proteins are extended (unlike that in the other SUTs), and this extended section may be involved in effector

binding or signal transduction (Barker at al., 2000). However, there is no direct evidence to date to support their role as sugar sensors. Some SUTs are also known to have substrate specificity for maltose (Riesmeier et al., 1992; Schulze et al., 2000; Kuhn, 2011). The monocot SUTs such as SUT1s of barley and sugarcane show more substrate specificity for maltose as compared to the dicot SUTs (Sivitz et al., 2005; Kuhn, 2011). For example, Chandran et al. (2003) found that SUT2 of Arabidopsis has low substrate specificity for maltose and other glucosides including arbutin, salicin, phenyl-a-glucoside, phenyl-b-glucoside, p-nitrophenyl-a-glucoside, p-nitrophenyl-b-glucoside, p-nitrophenyl-b-thioglucoside, turanose and a-methylglucoside. Similarly, SUT9 of Arabidopsis can transport glucosides such as helicin, salicin, arbutin, maltose, fraxin, esculin, turanose and a-methyl-D-glucose (Kuhn, 2011).

2.6.5 SUTs in cereal crops

Many *SUT*s have been identified from cereal crops including rice, barley, maize and wheat. Five *SUTs* are known in rice (Hirose et al., 1997) and two in barley (Weschke et al., 2000). The recently completed sequencing of the maize genome has revealed the presence of five maize *SUTs*. However, in hexaploid wheat, only one *SUT* (*TaSUT1*) has been identified to date. This gene has three homeologues designated as *TaSUT1A*, *TaSUT1B* and *TaSUT1D*, where A, B and D refer to the three genomes of hexaploid wheat (Aoki et al., 2002). The cDNA sequences of all the three homeologous genes show 96% identity among each other with differences at the 5' and 3'- untranslated (UTR) regions. Their predicted amino acid sequences exhibit 98% similarity (Aoki et al., 2002). The SUT1 proteins from rice, maize, wheat and barley show over 80% similarity in their

amino acid sequences, whereas the peptide sequence of SUT2 from barley (HvSUT2) shows 42.2% similarity with that of HvSUT1 (Weschke et al., 2000).

2.7 Role of cereal SUTs

Sucrose transported from the source tissues is released from the phloem into maternal tissues such as maize pedicel, and wheat, rice and barley pericarps. Since the maternal and filial tissues are separated by the placenta, chalazal or nucellar epidermal layers, sucrose uptake in the filial tissue is restricted. Therefore, apoplastic transport of sucrose is essential for starch accumulation to occur in the endosperm of cereal crops (Lemoine, 2000; Aoki et al., 2004).

2.7.1 SUTs in cereals are mainly involved in grain filling

In dicot plants, the role of SUTs in apoplasmic phloem loading is well studied, however, in monocots their role in grain growth and filling is the major focus of research. In cereal crops, the expression and localization patterns of *SUT* genes suggest that they are mainly involved in sucrose import from the grain apoplasm to the endosperm tissue. After its translocation through the phloem, sucrose is unloaded along the vascular strand situated in the crease of the grain. It then moves symplasmically from the sieve tubes to the nucellar projection transfer cells from where it is released into the endosperm cavity. The modified aleurone or sub aleurone layer separates the endosperm cavity from the starchy endosperm; therefore active uptake of sucrose, which is believed to be mediated by SUT proteins, is necessary for its transport into the endosperm (Aoki et al., 2004; Scofield et al., 2007).

Consistently, the SUT1 proteins of rice and barley are located in the plasma membrane of the aleurone layer, which surrounds the starchy endosperm (Weschke et al., 2000; Scofield et al., 2002). The gene encoding the SUT1 proteins in rice is expressed at very low level in the leaf blades after heading, but exhibits high expression in the developing endosperm (Hirose et al., 1997; Furbank et al., 2001; Sivitz et al., 2005). The barley *SUT1* (*HvSUT1)* is also found to be expressed in the caryopses endospermal transfer layer, which is the site of sucrose transport between the maternal and filial generation, suggesting a role in grain filling. *HvSUT1* expression is highest at 7-11 days after flowering. However, a general housekeeping role has been assigned to the barley *SUT2* (*HvSUT2),* as it is expressed in all sink and source tissues (Weschke et al., 2000). Like other cereal crops, in wheat *SUT1* transcripts accumulate highly in developing wheat grains (Aoki et al., 2004), and found to be localized in the maternal nucellar projection and the aleurone layer that separates the endosperm cavity from endosperm. Thus, active uptake of sucrose across the aleurone layer takes place with the assistance of SUTs.

Two stages of rice caryopsis development have been reported by Hirose et al. (2002): an early stage that encompasses elongation of caryopsis and cell differentiation (1 to 4 days of flowering) and the second stage that involves an increase in weight that starts from five days after flowering and lasts until 15 days after flowering, when the caryopsis has reached its full length. *OsSUT1* is expressed during the second stage i.e. grain filling stage and *OsSUT2, 3, 4* and *5* are shown to be mainly expressed during the first stage which suggests their role in caryopsis growth (Scofield et al., 2002; Aoki et al., 2003). This result shows the importance of SUTs for grain growth and filling in cereal crops.

Convincing evidence for the role of cereal SUTs in grain growth and filling is emerging. For example, the antisense repression of *OsSUT1* resulted in no effect on photosynthate supply from source leaves but caused substantial reduction in grain filling and eventual endosperm abortion (Ishimaru et al., 2001; Scofield et al., 2002).

2.7.2 Cereal SUTs mediate phloem loading

There is also evidence for the role of SUT proteins in phloem loading. The *SUT1* of maize (*ZmSUT1*) is highly expressed in green leaf blades and pedicels as compared to the kernels, and its expression in the source leaf blades was found to increase in parallel with the level of photoassimilates, indicating the role of *ZmSUT1* in phloem loading (Aoki et al., 1999). The accumulation of carbohydrates observed in the mature leaves of the *sut1* mutant of maize clearly demonstrated the role of *ZmSUT1* in phloem loading of sucrose (Slewinski et al., 2009). The inhibition of phloem loading in the mutant led to chlorosis that resulted in advanced senescence, decreased plant growth, and impaired reproductive development. Similar to that observed in maize, *SUT1* of wheat is highly expressed in flag leaf blades, flag leaf sheaths and culms at the pre heading stage, which may suggest that the TaSUT1 proteins play a phloem loading role as mesophyll cells and vascular cells show no symplastic connections (Aoki et al., 2004).

2.7.3 Cereal SUTs are involved in sucrose retrieval

Retrieval of sucrose leaked to the phloem apoplasm during transport is another function assigned to SUTs present in leaf sheaths and culms to ensure efficient transport of photoassimilates (Aoki et al., 2006). It has been shown that wheat SUT1 (TaSUT1)

protein is present in the phloem of leaf sheaths and internodes, suggesting its involvement in retrieving sucrose diffused into the apoplasm (Aoki et al., 2004). Similarly, SUT1 protein of rice (OsSUT1) is localized along the transport phloem of the leaf sheath and peduncle where it might be responsible for retrieval of sucrose and loading it back into the phloem (Scofield et al., 2007). The *SUT1* of barley (*HvSUT1)* is also expressed in maternal tissues i.e. the transfer cells of nucellar projections and epidermis where it may be involved in sucrose retrieval in maternal tissue (Weschke et al., 2000; Braun and Slewinski, 2009).

2.8 Functional characterization of SUTs with yeast

The wild type strain of yeast (*Saccharomyces cerevisiae*) is capable of growing on sucrose media due to the presence of extracellular hexose transporters and invertases, by which it is able to uptake and metabolize external sucrose (Riesmeir et al., 1992; Riesmeir et al., 1993; Lalonde et al., 1999). A mutant yeast strain, SUSY7, has been generated from the wild type yeast and is being used as tool for functional characterization of plant SUT proteins. In the SUSY7 strain, extracellular and cytosolic invertases are knocked out, then a plant sucrose synthase gene is expressed in the cytosol to make the mutant strain able to metabolize internal sucrose (instead of the external one) and thereby grow on sucrose containing media. The expression of a functional SUT protein on the plasma membrane of the mutant strain can internalize sucrose into the yeast cells, rendering them capable of growing on sucrose media as a sole carbon source, and thereby define the functionality of specific SUTs (Riesmeir et al., 1992; Riesmeir et al., 1993; Lalonde et al., 1999).

The mutant SUSY7 strain has been further modified through deletion of the URA3 gene (designated as SUSY7/ura3), making it incapable of producing uracil. This modified yeast strain can be transformed using vectors such as pDR196 that express URA3, which can be used as a selection marker for positive colonies when grown in a media lacking uracil (Barker et al., 2000). Previous studies have shown that SUSY7 yeast cells transformed with functional plant SUTs such as LeSUT1 (Barker et al., 2000), OsSUT1, OsSUT3 (Aoki et al., 2003) and HvSUT1 (Weschke et al., 2000), TaSUT1 (Aoki et al., 2006) are able to grow on sucrose containing media. With respect to the wheat SUT1 protein, the capabilities of all three TaSUT1 proteins in transporting sucrose and maltose were demonstrated by a complementation test with SUSY/ura3 mutant yeast cells. However, the capacity of these proteins for transporting maltose is less than that for sucrose (Aoki et al., 2006). Another yeast strain known as SEY6210 has been in use for determining the kinetic properties of sucrose transporter proteins. This strain is widely used for uptake experiments as it lacks not only invertase activity (like that of SUSY7) but also sucrose synthase activity in the cytosol, making it incapable of hydrolyzing internal sucrose. As a result, when the SEY6210 strain is transformed with a functional SUT protein, the sucrose transported by the expressed SUT will accumulate within the cells, making it an ideal system for quantitative analysis of sucrose uptake (Schulze et al., 2000).

2.9 SUT proteins as a tool for enhancing starch yield

Population growth along with increased diversified uses of cereal crops, for example, for bioethanol production, is leading to growing demands for these crops. This calls for the

need to improve the productivity of these crops. The grain starch yield in cereal crops is partly determined by the amount of sucrose transported from the source tissues to the endosperm during the grain filling phase. This in turn depends on the activities of SUTs, as the filial and maternal tissues of cereal grains lack symplastic connections (Weschke et al., 2000). Manipulation of the activity of SUTs would therefore result in increased grain starch content, and thereby enhanced crop yield. Improvement in crop yield by enhancing the activity of SUTs has been demonstrated. For example, expression of spinach SUT1 in potato resulted in reduction of sucrose level in the leaves and an increase in sugar levels in the tubers. Moreover, the plasma-membrane vesicles isolated from these plants exhibited an increased sucrose uptake rate as compared to those isolated from their wild type counterparts (Leggewie et al., 2003). The other evidence showing the significance of SUTs in enhancing yield come from transgenic pea plants that expressed potato SUT1 (a high affinity sucrose transporter) and thereby exhibited increased sucrose uptake capacity in the cotyledons, where the *StSUT1* gene was selectively expressed (Rosche et al., 2002). These results indicate that SUT proteins can play an important role in increasing the partitioning of photoassimilates towards the harvestable organs (Gifford et al., 1984). Therefore, identifying new SUTs and characterizing their functionality will have a significant contribution to designing tools for improved crop yield and quality.

3.0 Molecular Cloning and Functional Characterization of a New Sucrose Transporter in Hexaploid wheat (*Triticum aestivum* L.)

Abstract: Sucrose transporters play important roles in phloem loading and unloading, as well as retrieval of sucrose leaked out of the phloem during the transport process. This study reports the cloning and functional characterization of three homeologues of a sucrose transporter gene, from hexaploid wheat (*Triticum aestivum* L.) designated as *TaSUT2A*, *TaSUT2B* and *TaSUT2D*. The *TaSUT2A* and *TaSUT2B* genes each encode a protein with 506 amino acids whereas *TaSUT2D* encodes a 508 amino acid protein. The molecular mass of these proteins is predicted to be ~ 54 kDA. Phylogenetic and topological analysis of the putatitive amino acid sequences of the three genes reveals that they are new members of the GPH cation symporter family, each with 12 transmembrane helices, a characteristic feature of all known SUTs, as well as the histidine residue that is responsible for binding of sucrose during its transport. The ability of yeast SUSY7/*ura3* strain cells transformed with *TaSUT2A*, *TaSUT2B* and *TaSUT2D* to grow on a medium containing sucrose as the sole source of carbon clearly demonstrates their functionality as a sucrose transporter proteins. The presence of *TaSUT2* transcripts in the flag leaf blade, flag leaf sheath, peduncle, glumes, palea, lemma and grain tissues suggests its multi-functionality in phloem loading, unloading and sucrose retrieval.

3.1 Introduction:

Sucrose, the major transported form of carbohydrates, is exported from source to sink tissues via the phloem (Patrick and Offler, 1995) that contains sieve elements and companion cells as its key components. The sieve elements and companion cells of the phloem are connected by numerous intercellular connections called plasmodesmata, and they, together form the sieve element-companion cell complex (SE-CC). The sieve elements are mainly responsible for long distance transport of sugars and other organic materials whereas the companion cells act as a supplier of proteins and energy to the sieve elements (Williams et al., 2000; Truernit, 2001). Phloem loading and unloading of sucrose follows two pathways: symplastic, which involves cell to cell movement of sucrose via plasmodesmatal connections, and apoplastic, which involves active transport of sucrose across membranes when there are no plasmodesmatal connections between the cells (Truernit, 2001; Sauer, 2007). Thus, in apoplastic phloem loading, sucrose synthesized in the mesophyll cells of source leaves may enter the phloem apoplasm and be actively loaded into the SE-CCC with the help of plasma membrane localized sucrose transporter (SUT) proteins (Truernit, 2001; Sauer, 2007).

Starch synthesis in cereal grains is mainly dependent on the partitioning of photoassimilates produced in the source organs, and SUT proteins play critical roles with this respect. Genes encoding SUTs have been identified from a number of cereal crops. Rice SUT1 was the first cereal SUT to be identified (Hirose et al., 1997). In cereal crops such as wheat, barley and maize, the maternal and filial tissues lack symplastic connections (Weschke et al., 2000; Aoki et al., 2004; Lim et al., 2006), therefore, apoplastic transport of sucrose with the help of SUTs is essential for starch synthesis and

accumulation in the grains. Given that cereal crops are the most economically important and SUT proteins play significant roles during grain development, identification and characterization of *SUT* genes in cereals have recently been of primary attention (Kuhn, 2011).

Previous studies have shown that the SUT1s of barley (*HvSUT1*), wheat (*TaSUT1*) and rice (*OsSUT1*) are localized in the maternal nucellar projection and aleurone tissues, which separate the endosperm cavity from the endosperm, indicating the role of these genes in grain development (Weschke et al., 2000; Aoki et al., 2002; Lim et al., 2006). Consistent with localization studies, the expression of *OsSUT1* was found to be higher in the maternal nucellar projection and filial aleurone tissues as compared to that in the leaf blades, where its expression was low and detected only after heading (Hirose et al., 1997; Furbank et al., 2001; Sivitz et al., 2005). Thus, expression studies of *SUT1* from barley, rice, and wheat suggest that they play important roles in the transport of sucrose from maternal tissues, such as nucellar projection cells, and in its active uptake in the aleurone layer (Aoki et al., 2006; Lim et al., 2006). Given that the SUTs are mainly involved in post phloem sucrose transport, they are believed to play a significant role during grain filling in cereal crops (Aoki et al., 2002).

Expression analysis of barley *SUT* genes during grain development revealed that *HvSUT1* is highly expressed between 7 and 11 days after flowering (DAF), whereas the *HvSUT2* is expressed at 0 to 6 DAF and later at 18 to 21 DAF (Weschke et al., 2000). This indicates that *HvSUT1* is active during the rapid grain filling stage while *HvSUT2* is active during both initial and later stages of grain development. Since the SUT proteins also act as phloem loaders, they are also studied in source tissues. The *SUT2* genes from

barley (*HvSUT2*) and rice (*OsSUT2*) have been shown to have equal expression in grains as well as in source tissues such as leaves, suggesting that these genes have a housekeeping role (Weschke et al., 2000; Lim et al., 2006). Based on their sequence similarity, the *HvSUT2* and *OsSUT2* genes are classified into the SUT4 group which contains genes from both dicots and monocots (Braun and Slewinski, 2009; Kuhn, 2011, Aoki et al., 2012). These genes appeared to have very low sequence similarity with members of the monocot SUT1 group (Aoki et al., 2003). Members of the SUT4 group including *HvSUT2* and *OsSUT2*, as well as *SUT4* genes from Lotus and Arabidopsis have recently been found to be localized on the vacuolar tonoplasts (Endler et al., 2006; Kuhn and Grof, 2010), suggesting their involvement in vacuolar sucrose storage and transport (Kuhn, 2011). Knockout mutation of *SUT2* in rice resulted in reduction in plant height, tiller number and grain weight along with accumulation of sucrose in source leaves. All of these effects are attributed to defects in sucrose transport from vacuole to cytosol (Eom et al., 2011).

In Maize, *SUT1* (*ZmSUT1*) was found to be highly expressed in leaf blades and pedicels, as compared to grains (Aoki et al., 2002). Its expression in the leaf blades increased as the level of photoassimilates increased, reflecting its role in phloem loading (Aoki et al., 1999). The importance of *ZmSUT1* in phloem loading has recently been proven by a study that involved the *zmsut1* mutant, in which the impaired phloem loading resulted in accumulation of carbohydrates in mature leaves, leading to leaf chlorosis and senescence, and reduced plant growth (Slewinski et al., 2009). Expression of three wheat *SUT1* homeologues (*TaSUT1A*, *TaSUT1B* and *TaSUT1D*) was also detected in all source tissues and found to be equally expressed in the flag leaf, leaf sheaths and internodes,

suggesting a role in phloem loading. For *TaSUT1*, the highest transcript abundance in the source tissues was observed at 4 days before heading but showed a decrease by 12 days after heading (DAH), after which its transcript level in developing grains increased with the highest level detected by 16 to 20 DAH (Aoki et al., 2002). Supportive evidence for the role *TaSUT1* in phloem loading has come from a study that used a symplastic fluorescein tracer and showed that the SE-CC complex of the wheat flag leaf lacks plasmodesmatal connections (Aoki et al., 2004). Lack of symplasmic continuity has also been observed in the SE-CC complex of maize and barley source leaves (Botha and Cross, 1997; Haupt et al., 2001; Aoki et al., 2004). Furthermore, the presence of *TaSUT1* transcripts in the internodes suggests its role in the retrieval of sucrose that leaked out into the phloem apoplasm during transport, rendering efficient mechanism of photoassimilate translocation (Aoki et al., 2004). Unlike that observed in the leaves of wheat, maize and barley, the leaves of rice contain numerous symplastic connections between parenchyma and companion cells (Aoki et al., 2004). Moreover, the *SUT1* gene of rice is expressed at a very low level in source leaves (Hirose et al., 1997; Aoki et al., 2003; Aoki et al., 2004), and the antisense expression of *OsSUT1* resulted in impaired grain filling and retarded germination with no effect on photosynthesis in the flag leaf (Scofield et al., 2002). Though these results suggest that symplastic phloem loading operates in rice leaves, it has not yet been clearly defined which phloem loading mechanism is predominant (Lim et al., 2006).

In cereal crops such as wheat, the flag leaf is connected to the culm node through the leaf sheath (Aoki et al., 2004). The excess carbohydrates produced during the early stages of grain fiilling are temporarily stored in leaf sheaths and internodes in the form of

soluble carbohydrates, mainly as fructans, and are utilized during the later stages of grain filling, when photoassimilate supply decreases and sink demand increases (Schnyder, 1993; Haupt et al., 2001; Aoki et al., 2004). It has been indicated that approximately 50% of the photoassimilates that contribute to grain filling in wheat are temporarily stored in the stem and leaf sheaths prior to their remobilization during the active grain filling stage (Schnyder, 1993; Scofield et al., 2007). The detection of *OsSUT1* transcripts in the leaf sheath and stem of rice (Hirose et al., 1997; Furbank et al., 2001; Lim et al., 2006) suggests their role for remobilization of stored carbohydrates (Aoki et al., 2003). Beside these tissues, the palea, glume and lemma also have photosynthetic ability and therefore play a vital role in assimilates transport to the developing grain. Kreidemann (1996) and Wang et al. (2001) reported that the wheat ear photosynthetic capacity contributes about 10 to 44% of assimilates to the wheat grain depending on the type of genotypes and environmental conditions.

In the present study, we identified a new sucrose transporter gene, *TaSUT2*, from hexaploid wheat, characterized its expression in both source and sink tissues and examined its functionality using a yeast heterologous system. Moreover, the expression of *TaSUT2* in the three genomes that contributed to the hexaploid nature of common wheat and the phylogenetic relationship of the protein encoded by this gene with other known SUTs from dicots and monocots was examined.

3.2 Materials and methods

3.2.1 Plant growth conditions and tissue collection

Wheat plants (*Triticum aestivum*) cv. AC Andrew and the three diploid progenitors of hexaploid wheat, *Triticum urartu* (donor of A genome; accession # CN38564*), Aegilops speltoides* (donor of B genome; accession # CN108020) and *Aegilops tauschii* (donor of D genome; accession # PI560538) were used for this study. Mature dry seeds (kindly provided by Dr. Anita Brule Babel's Research Group) were imbibed on a moist sterile Whatman #1 filter paper placed in a Petri plate (15 seeds per plate) in darkness for three days prior to planting. Germinated seedlings were planted in 1-gallon pots (1 seedling per pot containing Sunshine Mix #4 (LA4; Sungro Horticulture, Bellevue, WA, USA,) and 18.7 g of nutrient supplement (Cornell mixture; 100 g calcium carbonate, 150 g osmocote, 120 g superphosphate, 2 g fritted trace elements, 1.5 g chelated iron [13.2%], 0.7 g chelated zinc [14%]) at an approximate depth of 2 cm. The pots were then placed in a growth chamber at 18/14^{o}C (day/night) under a 16/8 h photoperiod with cool white fluorescent light (175 µE $m^{-2}s^{-1}$) until harvest. The plants were watered every second day and fertilized with N-P-K (20:20:20) mix once a week.

For cloning the cDNAs of *TaSUT2*, young leaves harvested from 15 to 35 d-old plants of wheat cv. AC Andrew and the three diploid progenitors were used for RNA extraction. For all other studies including analysis of the spatio-temporal expression of *TaSUT2* and grain growth measurements, respective tissues were harvested at heading (8 days before anthesis; -8 DAA) and after anthesis (designated hereafter as DAA). To collect tissues at different stages after anthesis, plants were tagged when the first extrusion of the yellow anthers was evident. Flag leaf blade, flag leaf sheath and peduncle

tissues were harvested at heading and also at 5, 10, 15, 20, 25 and 30 DAA along with developing spikes from individual primary or secondary tillers (one tiller per plant per replication; 3 replications). Components of the spike including the glumes, lemmas, paleas and grains were separated from the middle region of each spike to minimize variations between samples. The tissues were immediately frozen in liquid nitrogen and stored at -80^{o} C until further usage.

3.2.2 Fresh and dry weight determination of grains

To determine the average fresh weight of developing grains at 5, 10, 15, 20, 25 and 30 DAA, 20 grains were weighed individually at each stage. Grain dry weight data was collected following incubation of the same grain samples in an oven at 105°C for 38 h.

3.2.3 RNA extraction

Total RNA samples from the flag leaf blade, flag leaf sheath, peduncle, glume, lemma and palea tissues were extracted using RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The frozen leaf tissues (~100 mg fresh weight) were ground to a fine powder in liquid nitrogen using a pre-chilled pestle mortar, after which 450 µL guanidinium thiocynate (Buffer RLT) containing 2-mercaptoethanol was added and mixed thoroughly. The mixture was then transferred to a QIAshredder spin column, centrifuged at 13,000*g* for 2 min, and the supernatant was separated and mixed with 100% ethanol (1/2 volume). The supernatant-ethanol mixture was transferred to RNeasy spin column and centrifuged at 8000*g* for 15 s, after which 700 µL of buffer RW1 (containing guanidinium thiocynate and ethanol) was added to the spin column

followed by centrifugation at 8000*g* for 15 s. Buffer RPE (containing 100% ethanol; 500 µL) was added to the spin column and centrifuged at 8000*g* for 15 s. Washing the spin column with ethanol was repeated except that this time the centrifugation was extended for 2 min to remove any ethanol remnant. Finally, the RNA was eluted from the spin column by adding 30 µL of RNase-free water and subsequent centrifugation at 8000*g* for 1 min.

Total RNA samples from developing grains were isolated according to Li and Trick (2005). Grains (~100 mg fresh weight) were ground to a fine powder as described above. The ground tissues were mixed thoroughly with 400 µL extraction buffer I (containing 100 mM Tris, pH 8, 150 mM LiCl, 50 mM EDTA, 1.5% sodium dodecyl sulfate and 1.5% 2-mercaptoethanol). After initial extraction with buffer I, 250 µL of phenol/chloroform/isoamyl alcohol (PCI at 125:24:1 ratio, pH 4.7) mixture was added to the samples and mixed by inverting the tubes, followed by centrifugation at 13,000*g* at 4°C for 15 min. The upper aqueous layer (~ 250 µL) was transferred to a new RNase free 1.5 mL tube and precipitated by using 250 µL buffer II (containing 4.2 M guanidinium sulfate (w/v), 25 mM sodium citrate, 0.5% lauryl sarcosine, and 1 M sodium acetate, pH 4) followed by mixing with gentle inversion and incubation at room temperature for 10 min. A PCI mixture (200 µL) was added to the sample and then centrifuged at 13,000*g* at 4°C for 15 min. The resulting supernatant was further precipitated with 300 µL of isopropanol and 250 µL of 1.2 M sodium chloride. After mixing with inversion, the mixture was incubated on ice for 15 min, and then centrifuged for 15 min at 13,000*g* at 4^{o}C. The resulting pellet was washed carefully with 70% ethanol, allowed to dry for 15-

20 min at room temperature and then re-suspended in 50 µL diethylpyrocarbonate (DEPC) treated water.

The integrity of the resulting RNA was verified by gel electrophoresis, and the purity by the average spectrophotometric absorption ratio at 260 to 280 nm and 260 to 230 nm. To eliminate genomic DNA contamination, the total RNA samples were digested with DNAase (DNA-free Kit; Ambion, Austin, TX, USA). Sample RNA concentration was determined in duplicate by A_{260} measurement, and then the samples were stored at -80^{o}C until cDNA synthesis.

3.2.4 cDNA synthesis

Complimentary DNA (cDNA) was synthesized from total RNA using RevertAid™ H Minus First Strand cDNA Synthesis Kit (Fermentas, Glen Burnie, MD, USA) following the manufacturer's protocol. Total RNA (1 µg) was mixed with 1 µL of 100 µM oligo (dT) primer in a total volume of 12 µL, and incubated at 65^{o}C for 5 min. After incubation on ice, 4 µL of 5X reaction buffer, 2 µL of 10 mM deoxyribonucleotide triphosphate (dNTP), 1 µL RiboLock RNase Inhibitor (20 U/ µL) and 1 µL of RevertAid M-MuLV Reverse Transcriptase (200U/µL), were added, resulting a total reaction volume of 20 µL. The mixture was incubated at 42^{o}C for 1 h, and then the reaction stopped by incubating the mixture at 70^{o}C for 10 min. The concentration of the cDNA was determined by Nanodrop (ND-1000, ThermoFisher Scientific; Waltham, MA, USA) and then stored at -20^{o}C until used for partial and full length amplification of TaSUT2. First strand cDNA synthesis for gene expression analysis was performed using 1 µL of 100

µM random hexamer primers [instead of oligo (dT) primer],which involved additional incubation at 25°C for 5 min prior to incubation at 42°C for 1 h.

3.2.5 Isolation of partial sequence of *TaSUT2*

3.2.5.1 PCR amplification

In order to amplify the partial sequence of *TaSUT2* specific forward (TaSUT2P-FW) and reverse (TaSUT2P-R) primers (Table 3.1) were designed (using DNAman software) from the conserved coding sequences of reported *SUT2* genes from barley (*HvSUT2*, GenBank accession number AJ272308; Weschke et al., 2000), maize (*ZmSUT2*, GenBank accession number AY639018.1; Wright et al., 2005) and rice (*OsSUT2*, GenBank accession number AB091672; Aoki et al., 2003) (Appendix 1). Amplification of the partial sequence was performed by using thermostable Phusion High-Fidelity DNA Polymerase (Finnzymes, Vantaa, Finland). Each 50 µL PCR reaction contained cDNA (~50 ng) as a template, 10 µL of 5X Phusion High-Fidelity buffer, 1 µL of 10 mM dNTPs, 2 µL of 10 µM forward primer (400 nM final concentration), 2 µL of 10 µM reverse primer (400 nM final concentration), 0.5 µL Phusion DNA polymerase (2 U/µL) and 32 µL sterile distilled water. The PCR reaction mixtures were subjected to thermal cycling conditions of initial denaturation at 98^{o}C for 30 s followed by 35 cycles of 98^{o}C for 10 s, 60^{o}C for 30 s and 72^{o}C for 50 s followed by final extension at 72^{o}C for 7 min using a DNA thermal cycler (Bio-Rad; Herculus, CA, USA). Following amplification, samples were loaded into a 1% (w/v) agarose gel, electrophoresed, and then stained with ethidium bromide. The size of the expected PCR product was determined by comparing with a DNA marker (1Kb+).

Table 3.1 Primer sequences used and their respective amplicon sizes.

Name	Purpose	Sequence	Amplicon
TaSUT2P-FW[a] TaSUT2P-R[b]	Isolation of *TaSUT2* partial length	5'-CATGGCCCTGGGGAACATA-3' 5'-CAGCGACACGATGATCTGTGG-3'	808 bp[c]
TaSUT2F-FW TaSUT2F-R	Isolation of *TaSUT2* full length	5'-ACGCGGGGAGAGCGATAGATAC-3' 5'-TTTCAAACAAAAACCTGCAACTTTTATTG-3'	2036 bp
M13 FW M13R	Sequencing *TaSUT2* insert into pGEM-T	5'-GTAAAACGACGGCCAGT-3' 5'-GCGGATAACAATTTCACACAGG-3'	N/A
Spe1-TaSUT2 (FW) TaSUT2-*Xho1*(R)	*TaSUT2* insertion into pDR196	5'-TACTAGTATGCCGCCGCGGC-3' 5'-GTCTCGAGTCATCGTTGGGTGGTT-3'	2051 bp
EcoR1-TaSUT2 (FW) TaSUT2-*Xho1* (R)	*TaSUT2* insertion into pENTR1A	5'-AGGAATTCATGCCGCCGCGGC-'3 5'-CGCTCGAGTCATCGTTGGGTGGTT-'3	2052 bp
TaSUT2-FW1 TaSUT2-R1	qPCR	5'-TACGGAGTCCTGCTCTGTCA-3' 5'-CTCGTCGCTTCCGAAAGTA-3'	130 bp
Taβ-actin (FW) Taβ-actin (R)	qPCR	5'-CCTTCCACATGCCATCCTTC-3' 5'-GCTTCTCCTTGATGTCCCTTAC-5'	137 bp

[a] Forward Primer, [b] Reverse Primer, [c] base pair

3.2.5.2 Gel extraction

Following visualization with a UV transilluminator (Fisher Scientific, Hampton, NH, USA), a band of expected size was excised from the gel and the DNA fragment purified using GeneJET Gel Extraction Kit (Fermentas) according to the manufacturer's instructions. The excised DNA fragment was placed into a microcentrifuge tube to which the same volume of binding buffer (guanidinium thiocyanate) was added. The mixture was incubated at 50^{o}C until completely dissolved, and then one gel volume of isopropanol was added. The mixture was transferred onto the GeneJET purification column and centrifuged at 13,000*g* for 1 min. Binding buffer (100 μL) was added to the column in order to remove any residual gel and then centrifuged at 13,000*g* for 1 min

followed by washing with 700 µL of wash buffer (100% ethanol) and centrifugation for 1 min. The column was centrifuged further for 1 min to ensure complete removal of the wash buffer, and then transferred into a clean microcentrifuge tube for eluting the purified DNA using nuclease free water and centrifugation at 13,000*g* for 1 min. The purified DNA was stored at -20^{o}C until further use.

3.2.5.3 Ligation and transformation

The purified DNA was ligated into pGEM-T Easy Vector (Promega, Madison, WI, USA). The ligation reaction contained the purified DNA fragment (75 ng), pGEM-T Easy Vector (50 ng), 5 µ1 of 2X rapid ligation buffer, 1 µ1 of T4 DNA ligase (3 U/µL) and 1 µ1 sterile water, with a total reaction volume of 10 µ1. The reaction mixture was incubated overnight at 4^{o}C, and the ligation product was used to transform *E.coli* DH5α competent cells (Invitrogen, Carlsbad, CA, USA). Briefly, 5 µL of the ligation product (~ 10 ng) was mixed with competent DH5α cells (50 µL) followed by incubation on ice for 30 min and subsequent heat shock at 42^{o}C for 50 s. After incubation on ice for 2 min, 950 µL of Luria-Bertani (LB) media (containing 10 g tryptone, 5 g yeast extract and 10 g NaCL per liter; pH 7.0) was added and incubated on a shaker (New Brunswick Scientific, Edison, NJ, USA) at 37^{o}C for 1 h. The transformation mixture (200 µL) was plated on LB agar plates containing 100 µg of ampicillin (Sigma, St. Louis, MO, USA), 500 µM isopropyl β-D-1-thiogalactopyranoside (IPTG; Fermentas) and 80 µg X-gal (Promega) per 1 mL of LB. The plates were incubated at 37^{o}C overnight. The resulting white colonies were selected and screened by PCR using reaction mixtures and thermocycling conditions as described above.

3.2.5.4 Plasmid isolation

Plasmid extraction was performed using the GeneJET Plasmid Miniprep Kit (Fermentas). A 10 mL LB media was inoculated with PCR screened positive colonies followed by incubation on a shaker (250 rpm) at 37^{o}C overnight. The resulting culture was centrifuged at 3000*g* for 5 min and the supernatant was discarded. The pelleted cells were mixed with 250 µL of resuspension solution containing RNase A followed by addition of 250 µL of lysis solution containing sodium dodecyl sulphate (SDS) and sodium hydroxide to assist the liberation of plasmid DNA. Then 350 µL of neutralization solution containing guanidinium chloride and acetic acid (enhances the binding of plasmid DNA on the silica membrane of spin column) was added and mixed by inverting the tubes several times to avoid precipitation of bacterial cell debris. The mixture was then centrifuged at 13000*g* for 5 min to pellet cell debris and chromosomal DNA, and the resulting supernatant was transferred to GeneJET spin column and centrifuged at 13000*g* for 1 min. Following addition of 500 µL wash solution and centrifugation at 13000*g* for 1 min (2X), the spin column was transferred to a clean microcentrifuge tube and the plasmid DNA eluted with 50 µL of elution buffer. The purified plasmid was stored at -20^{o} C until further use.

3.2.5.5 DNA sequencing

The purified plasmid was sequenced (Macrogen, Rocville, MD, USA) using plasmid specific M13 forward and reverse primers (Table 3.1). In order to examine if the isolated DNA fragment represent a putative partial sequence of *TaSUT2* gene, its nucleotide

sequence was blasted against GenBank database using the Basic Local Alignment Search Tool (BLAST; Altschul et al., 1990).

3.2.6 Isolation of 3' and 5' end fragments of *TaSUT2*

3.2.6.1 First strand 3' and 5' RACE ready cDNA synthesis

The identified putative partial sequence of *TaSUT2* was used to design gene specific primers to isolate the 3' and 5' end fragments of *TaSUT2* with RACE (Rapid Amplification of cDNA Ends)-PCR using SMARTer RACE cDNA Amplification kit (Clonetech, Mountain View, CA, USA). To this end, first strand 3' and 5' -RACE ready cDNAs were prepared according to the manufacturer's instructions. In one tube, total RNA (1 μg) was mixed with 1 μL of 12 μM 3'-RACE CDS Primer A (Clontech) for synthesizing 3' RACE cDNA, and in another tube, total RNA (1 μg) was mixed with 1 μL of 12 μM 5'-RACE CDS Primer A for 5' RACE cDNA synthesis. These two reaction mixtures were incubated first at 72^{o}C for 3 min and then at 42^{o}C for 2 min before centrifugation at 14000*g* for 10 s. Then, 1 μL of 12 μM SMARTer IIA oligo was added only to the 5' RACE cDNA. A master mix containing 2 μL of 5X first strand buffer, 1 μL of 20 mM dithiothreitol (DTT), 1μL of 10 mM dNTPs, 0.25 μL RNase inhibitor (40 U/μL) and 1 μL of SmartScribe Reverse Transcriptase (100 U/μL) was prepared and then added to the tubes containing total RNA samples and RACE CDS Primer A. The resulting mixture was incubated first at 42^{o}C for 90 min and then at 70^{o}C for 10 min in a thermal cycler (Bio-Rad). Finally, the synthesized first strand 3' and 5' RACE Ready cDNA was diluted with Tricine-EDTA buffer and stored at -80^{o}C until further use.

3.2.6.2 Isolation of 3' end fragment

The first strand 3' RACE ready cDNA was used as a template for 3' RACE-PCR. The 3' end fragment of *TaSUT2* was amplified using thermostable *Taq* DNA polymerase (Fermentas), and outer gene specific (TaSUT2-3'OP2; Table 3.2) and universal (Clontech) primers with PCR reactions and thermocycling conditions described previously. The 3' outer RACE product was then used as a template for nested PCR, which was performed using inner gene specific (TaSUT2-3'NP2; Table 3.2) and nested universal (Clontech) primers. The products of both 3' outer and inner RACE PCR were separated by gel electrophoresis, purified, cloned into the pGEM-T Easy vector and then sequenced as described previously. The identity of the nucleotide sequence of the newly identified 3' end fragment was examined by BLAST analysis of its sequence against GenBank database using BLAST.

Table 3.2 Sequences of gene specific primers used for isolation of 5' and 3' end fragments of *TaSUT2*.

Name	Position	3' or 5' RACE	Sequence
TaSUT2-3'OP2	Outer	3' RACE	5'-CGCCGCTTCCCTCATTGTGTTCACA-3'
TaSUT2-3'NP2	Inner	3' RACE	5'-ATGGCTGCAAGTCGTGTTGAA-3'
TaSUT2-5'OP1	Outer	5' RACE	5'TCCGAAAGTAGGATTGTCCTGCACCGTC-3'
TaSUT2-5'NP1	Inner	5' RACE	5'-AGGTTGGCGCAGCTGACAGA-3'
TaSUT2-5'OP2	Outer	5'extension	5'-GACGTACGGGGTGAGCAGGGAGA-3'
TaSUT2-5'NP2	Inner	5'extension	5'-AGCGAGCGGAGCGGGACCTT-3'

3.2.6.3 Isolation of 5' end fragment

The first strand 5' RACE ready cDNA was used as a template for 5' RACE-PCR. The 5' end fragment of *TaSUT2* was amplified using thermostable Phusion High-Fidelity DNA polymerase (Finnzymes) using outer gene specific (TaSUT2-5'OP1; Table 3.2) and universal (Clontech) primers. The PCR reaction mixture prepared as described above was subjected to touchdown PCR using a thermocycling condition of 5 cycles (98^{o}C for 30 s and 72^{o}C for 1 min); followed by 5 cycles (98^{o}C for 30 s, 70^{o}C for 30 s and 72^{o}C for 1 min) and 25 cycles (98^{o}C for 30 s, 68°C for 30 s and 72^{o}C for 1 min) and final extension at 72^{o}C for 10 min. The 5'outer RACE PCR product was subjected to nested PCR using an inner gene specific primer (TaSUT2-5'NP1; Table 3.2) and the nested universal (Clontech) primers. The 5' inner RACE PCR product was excised from the gel, purified and cloned into the pGEM-T Easy vector, and then sequenced as described previously. Since translation of the identified 5' end fragment showed an incomplete protein domain, it was extended further by RACE-PCR using a nested set of gene specific primers (TaSUT2-5'OP2 and TaSUT2-5'NP2; Table 3.2) designed from the newly identified 5' end fragment and the universal (Clontech) primers.

3.2.6.4 End to end PCR

After isolation of the 3' and 5' end fragments of *TaSUT2* by RACE-PCR, these two fragments were assembled along with the already identified partial sequence to produce a putative full length cDNA sequence of *TaSUT2*. End to end PCR was performed by using forward (TaSUT2F-FW) and reverse (TaSUT2F-R) primers (Table 3.1) designed from the 5' and 3' ends of the putative full length sequence, respectively, and thermostable

Phusion High-Fidelity DNA polymerase (Finnzymes) as described previously. Following amplification, the PCR product was separated with 1% (w/v) agarose gel, and then the DNA band corresponding to the full length of *TaSUT2* was excised and purified. The purified fragment of *TaSUT2* was ligated into pGEM-T Easy vector, and *E.coli* (DH5α) cells were transformed with the ligated product. Following PCR screening of the positive colonies, plasmid isolation and subsequent sequencing of the *TaSUT2* was performed as described above.

3.2.6.5 Sequence analysis

The identity of *TaSUT2* was analyzed by blasting its nucleotide sequence against GenBank database using BLAST, and its coding sequence identified by using Open Reading Frame (ORF) finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). The coding nucleotide sequence was translated into amino acid sequence by using the JustBio translator tool (http://www.justbio.com/index.php?page=hosted-tools). The resulting amino acid sequence was blasted against GenBank database to compare its similarity with amino acid sequences of known SUT proteins. The molecular weight of TaSUT2 was determined using the protein molecular weight calculator software (http://www.sciencegateway.org/tools/proteinmw.htm). Its transmembrane topology was predicted by using TMpred (http://www.ch.embnet.org/software/TMPRED_form.html).

3.2.7 Identification of *TaSUT2* from the A, B and D genomes of wheat

To detect the presence of *TaSUT2* in the three (A, B and D) genomes of hexaploid wheat, cDNA samples prepared from each of the three diploid progenitors of hexaploid wheat

were first amplified using *TaSUT2* specific forward and reverse primers (Table 3.1). The amplified fragments of *TuSUT2, AesSUT2 and AetSUT2* were gel purified and then ligated into pGEM-T Easy vector, and the ligation product was used to transform *E.coli* (DH5α) cells. Following PCR screening of the positive colonies, plasmid isolation and subsequent sequencing of the cDNAs of *SUT2* from the progenitors was performed as described above. The presence of the three homeologues of *TaSUT2* (*TaSUT2A, TaSUT2B, TaSUT2D*) in hexaploid wheat cv. AC Andrew was examined by sequencing multiple colonies obtained by transformation of *E.coli* (DH5α) cells with *TaSUT2* identified by the RACE-PCR.

3.2.8 Phylogenetic relationship of TaSUT2 with other known SUTs

Protein sequences of known SUTs of both monocot and dicot species were collected from GenBank database and aligned with that of TaSUT2 using the ClustalW alignment tool from the Molecular Evolutionary Genetic Analysis (MEGA, version 5) software (Tamura et al., 2011). An unrooted neighbor-joining phylogenetic tree was generated using MEGA5 software with a Poisson correction model and a 500 replicate bootstrap method of phylogeny test.

3.2.9 Recombinant protein expression:

Expression of the TaSUT2 protein was performed with *E.coli* expression system of Gateway Technology (Invitrogen). The open reading frame of *TaSUT2* was amplified using *EcoR1* site linked forward primer [*EcoR1*-TaSUT2 (FW)] and *Xho1* site linked reverse primer [TaSUT2-*Xho1* (R); Table 3.1]. The resulting PCR product was separated

by gel electrophoresis, excised, and purified as described previously. Both the purified *TaSUT2* fragment and the pENTR1A entry vector were digested with the *EcoR1* and *Xho1* restriction enzymes, which do not cut the *TaSUT2*. The digested products of *TaSUT2* and pENTR1A and were purified and subsequently ligated, and the ligation product was used to transform DH5α *E.coli* cells. The positive clones containing pENTR1A-*TaSUT2* were identified by PCR screening. Following isolation of plasmids from the positive colonies, *TaSUT2* was sub-cloned into expression vector pDEST14 using recombination reaction and LR clonase mix according to manufacturer's instructions. The 10 µL recombination reaction contained 100 ng of plasmid (pENTR1A-*TaSUT2*), 150 ng of pDEST14, 5 µL of TE buffer (pH 8) and 2 µL of LR clonase. The reaction mixture was incubated at 22°C for 2 h followed by addition of 1 µL protease K and further incubation at 37°C for 10 min. The LR reaction product was transformed with *E.coli* (BL21-DE3) cells (Novagen, Darmstadt, Germany) using the same protocol as for transforming the DH5α cells.

Transformed positive colonies were grown in 100 mL of LB media at 37^{o}C overnight, and the cell culture was then used to inoculate 500 ml of fresh LB medium containing 100 mg/L of ampicillin and allowed to grow till an optical density of 0.6 was achieved at 600 nm wavelength (OD_{600}). The resulting culture was then divided into two parts, and 250 µL of 1M IPTG was added (final concentration of 1 mM) to one part of the culture to induce the expression of TaSUT2. Both cultures were then incubated further for 3 h at 37°C, and their cells were harvested by centrifugation at 5000*g* for 10 min at 4^{o}C and then resuspended with 1 ml of lysis buffer (containing 50mM of Tris-Cl, pH 7.5). Following disruption by sonication, the cells were centrifuged at 13,000*g* for 15 min

at 4^{o}C. The supernatant containing the crude protein sample was analyzed by sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE).

3.2.10 Functional Analysis of TaSUT2

3.2.10.1. Cloning *TaSUT2* into expression vector

Cloning of *TaSUT2* into an expression vector involved amplification of its ORF using restriction site linked primers and thermostable Phusion High-Fidelity DNA polymerase. The *Spe1* restriction site was linked with the forward primer [*Spe1*-TaSUT2 (FW)] and *Xho1* restriction site was linked with the reverse primer [TaSUT2-*Xho1* (R); Table 3.1]. Restriction analysis of *TaSUT2* cDNA was performed to make sure that either of these enzymes do not cut *TaSUT2* gene. The resulting PCR product was separated by gel electrophoresis and purified as described above. Both the purified *TaSUT2* cDNA and pDR196 expression vector (kindly provided by Dr. John Ward's lab, Department of Plant Biology, University of Minnesota, USA) were then digested with FastDigest *Spe1* and *Xho1* restriction enzymes (Fermentas) prior to ligation. The total volume of 30 µL digestion reaction contained 1µg of plasmid DNA (for pDR196) or 200 ng (for *TaSUT2*), 2 µL of 10X Fast Digest buffer, 1 µL of *Spe1* (1 U/µL) and 1 µL of *Xho1* (1 U/µL) and 17 µL of water. The reaction mixtures were incubated at 37^{o}C for 20 min in a thermal cycler (Bio-Rad), after which the digestion products were separated by gel electrophoresis. The purified digested products of *TaSUT2* and pDR196 were ligated using a 10 µL total reaction mix that contained 5 µL of 2X ligation buffer, 50 ng of digested pDR196, 75 ng of digested *TaSUT2* and 1 µL T4 DNA ligase (3 U/µL). The ligation reaction mixture was incubated at 4^{o}C overnight and then used to transform

E.coli (DH5α) cells as described earlier. White colonies were selected and then PCR screened to verify that they contain the *TaSUT2* insert. Plasmid DNA extracted from the positive colonies was used for sequencing. Furthermore, the isolated plasmid was digested with both *Spe1* and *Xho1* and then electrophoresed for verifying the the insert size.

3.2.10.2 Preparation of yeast competent cells

The functionality of *TaSUT2* in transporting sucrose was examined by transforming competent cells of yeast *(Saccharomyces cerevisiae*) SUSY7/*ura3* strain (kindly provided by Dr. John Ward's lab) with the pDR196-TaSUT2 construct. For preparing yeast competent cells, 10 µL glycerol stock of SUSY7/*ura3* strain was mixed with 5 mL of yeast peptone dextrose (YPD; containing 50 g YPD broth mix/L) and incubated on a shaker (250 rpm) at 30^{o}C overnight. Then 100 mL of YPD media was inoculated with 5 mL of the overnight culture followed by incubation on a shaker (250 rpm) at 30^{o}C until an OD_{600} of 0.8 was achieved. Optical density was analyzed with Ultrospec 3100 pro spectrophotometer (Biochrom, Holliston, MA, USA). The cells were then pelleted by centrifugation at 1000*g* for 7 min followed by washing with 50 mL of Solution 1 (containing 1 M sorbitol, 10 mM Bicine-NaOH and 3% polyethylene glycol). Finally, the pellet containing the competent cells was resuspended in 2 mL of Solution 1, divided into 200 µL aliquots and stored at -80^{o}C.

3.2.10.3 Yeast transformation with pDR196-TaSUT2

To transform the yeast strain with pDR196-TaSUT2, 5 µg of pDR196-TaSUT2 was mixed with 200 µL of frozen suspension of competent yeast cells followed by the addition of 1.2 mL of Solution 2 (containing 0.2 M Bicine-NaOH and 40% polyethylene glycol, pH 8.4) and incubation at 30^{o}C for 1 h with no shaking. The cells were then spun at 5000*g* for 15 s and resuspended with 1 mL of Solution 3 (containing 150 mM sodium chloride, 10 mM bicine-NaOH, pH 8.35). Transformed SUSY7/*ura3* cells were grown at 30^{o}C on synthetic dextrose (SD) media with glucose as sole carbon source. The SD media contained 1.7 g/L yeast nitrogen base, 5 g/L ammonium sulphate, 20 g/L glucose, 10 mL/L of each of 100X L-leucine, L- histidine, L-lysine and L-tryptophan, and 20 g/L agar for plates. Transformed SUSY7/*ura3* cells were also grown on synthetic complete (SC) media with sucrose as sole carbon source. The SC media contained 1.7 g/L yeast nitrogen base, 5 g/L ammonium sulphate, 20 g/L sucrose, 10 mL/L of each of 100X L-leucine, L- histidine, L-lysine and L-tryptophan, and 20 g/L agar for plates. The yeast colonies transformed with pDR196-TaSUT2 were verified performing colony PCR using *TaSUT2* specific forward and reverse primer sequences linked with the restriction sites described above. The SUSY7/ura3 strains were also transformed with potato SUT1 (pDR196-StSUT1) and empty pDR196, as described above, to be used as positive and negative controls, respectively.

3.2.11 Real time PCR assay

Gene specific primers for *TaSUT2* and *TaActin* (used as a reference gene) were designed using Primer3 software (http://frodo.wi.mit.edu/primer3/; Table 3.1). Primer specificity

was confirmed by blasting the amplicons against GenBank database, and furthermore by separating RT-PCR products amplified with the specific primers using 1.5% agarose gel electrophoresis. Real time PCR assays were performed using Maxima® SYBR Green/ROX qPCR Master Mix (Fermentas). The reaction mixture contained 2 µL of cDNA (100 ng/µl), 10 µL of Maxima SYBR Green/ROX qPCR Master Mix, 0.6 µL of forward primer (10 µM; 300 nM final concentration), 0.6 µL of reverse primer (10 µM; 300 nM final concentration) and 6.8 µL of water, with a total reaction volume of 20 µL. Amplification and fluorescent signal detection was performed on a Mx3000P Real time PCR System (Stratagene, La Jolla, CA, USA) using the following thermocycling conditions: initial denaturation and DNA polymerase activation at 95°C for 10 min, followed by 40 cycles of denaturation at 95°C for 15 s, annealing and extension at 60°C for 1 min in 96-well optical reaction plates covered with optical caps (Bio-Rad). The relative transcript level of *TaSUT2* was determined by $2^{-\Delta\Delta Ct}$ (Livak and Schmittgen, 2001).

3.2.12 Statistical analysis

Statistical analysis of the gene expression data was performed by using the mixed procedure model of SAS (Ver. 9.2., SAS Institute Inc, 2008). The analysis of variance was performed to compare differences in expression of *TaSUT2* in developing grains and potential source tissues at different time points after anthesis using Tukey's LSD test ($P < 0.05$).

3.3 Results

3.3.1 Molecular cloning of *TaSUT2*

3.3.1.1 Isolation of partial sequence of *TaSUT2*

Amplification of the cDNA samples prepared from the young leaf tissues of wheat cv. AC Andrew with forward and reverse primers designed from the conserved region of *SUT2* genes previously identified from barley (*Hordeum vulagare*, *HVSUT2*), rice (*Oryza sativa*, *OsSUT2*) and maize (*Zea mays*, *ZmSUT2*; Appendix 1) produced a DNA fragment of the expected size (~808 bp; Figure 3.1A).

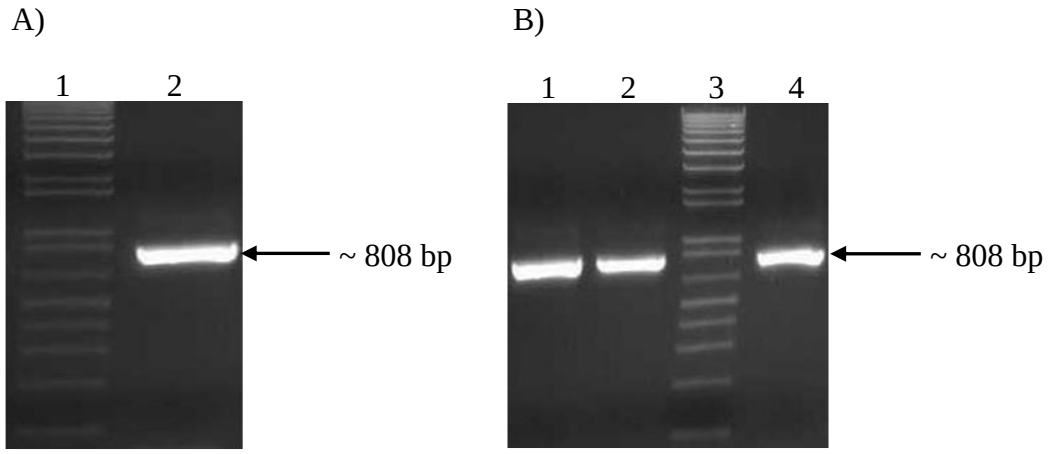


Figure 3.1 PCR product of the partial DNA fragment of *TaSUT2* amplified from cv. AC Andrew using primers designed from the conserved regions of *SUT2*s of other monocot species (A; lane 2). PCR screening of *E.coli* (DH5α) cells transformed with the partial DNA fragment of *TaSUT2*-pGEM-T construct (B); lanes 1, 2, 4, show DNA fragments of expected size (~808 bp) amplified from the positive colonies. Lanes 1 (A) and 3 (B) refer to the DNA marker (1Kb+) used to determine the size of PCR products.

Cloning of the amplified PCR fragment into pGEM-T Easy vector followed by transformation with *E.coli* (DH5α) cells produced colonies. Subsequent screening by PCR using primers designed from the conserved regions of the other *SUT2* genes

identified positive colonies containing the inserted DNA fragment (Figure 3.1B). Sequencing of the DNA fragment revealed that it has a size of 808 bp (Figure 3.2). A search for nucleotide sequence similarity was performed by blasting the identified putative *TaSUT2* partial sequence against GenBank database. The sequence of the newly identified DNA fragment of *TaSUT2* showed 98% identity with *SUT2* like sequence from wheat cv. Bobwhite (Table 3.3, Appendix 2). However, the *SUT2* like gene from cv. Bobwhite does not have a full length nucleotide sequence exhibited by members of the *SUT2* gene family and its protein lacks the domains that are characteristic to SUT2 proteins. The identified putative *TaSUT2* partial nucleotide sequence also showed 94% and 78% identity with *SUT2* genes from barley and rice, respectively (Table 3.3, Appendix 2).

CATGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGA
TATTCCCCTTCACTATTACGGAGTCCTGCTCTGTCAGCTGCGCCAACCTCAAGTCCGCG
TTCCTGCTCGACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGTGACGGTGCA
GGACAATCCTACTTTCGGAAGCGACGAGGCGGCACCTCCGAGCAGCCACGAGGAGGAGG
CTTTCCTCTTTGAGCTTTTCGGGTCGTTCAAGTACTTCACGATGCCTGTTTGGATGGTC
TTGATCGTCACCTCGCTTACCTGGATCGGCTGGTTCCCTTTCATCCTCTTTGACACCGA
CTGGATGGGCCGAGAGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAAAAGT
ATCATGACGGTGTGAGAATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCTCGGG
ATCACATCTATTGGAATGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTATGGGG
TGTCTCCAATATCATCATGGCACTGTGCTTCGTGGCGATGCTTATTATAACATACGTTG
CGCAGAACCTGGATTATGGACCTAGTGGAGCACCTCCAACCGGCATTGTCGCCGCTTCC
CTCATTGTGTTCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTATACCATATGC
GATGGCTGCAAGTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGCATTC
TTAATTTATCTATTGTCATACCACAGATCATCGTGTCGCTG

Figure 3.2 Nucleotide sequence of the partial DNA fragment of *TaSUT2* (808 bp) amplified from cv. AC Andrew. The forward and reverse primers used for amplifying the fragment are underlined.

Table 3.3 Nucleotide sequence identity between the partial DNA fragment of *TaSUT2* and *SUT2s* from other species.

Accession #	Description	Max Score[a]	E-Value[b]	Max Identity[c]
EU693244.1	*Triticum aestivum* cv. Bobwhite (*SUT2* like)	1437	0.0	98%
AJ272308.1	*Hordeum vulgare SUT2*	1232	0.0	94%
AB091672.1	*Oryza sativa SUT2*	525	2e-145	78%
GU045300.1	*Sorghum bicolor SUT4*	523	6e-145	78%
AY639018.1	*Zea mays SUT2*	501	3e-138	78%

[a]Maximum score indicates the amount of sequence homology between *TaSUT2* and *SUTs* from other species.
[b]E-Value determines the significance of the alignment score.
[C]Maximum identity refers to percentage of identical nucleotides between *TaSUT2* and the respective SUTs from other species.

3.3.1.2 Isolation of 5' and 3' end fragments of *TaSUT2*

The putative partial sequence of *TaSUT2* was used to design gene specific primers for isolation of 5' and 3' end fragments of *TaSUT2* by RACE-PCR. Amplification of the cDNA samples prepared from young leaf tissues of wheat cv. AC Andrew with gene specific primers designed from the 5' and 3' ends of the partial sequence (Table 3.2, Appendix 3, 4), and inner and outer universal primers (Clontech) produced 5' and 3' end DNA fragments with approximate size of 659 bp (Figure 3.3A) and 759 bp (Figure 3.3B), respectively.

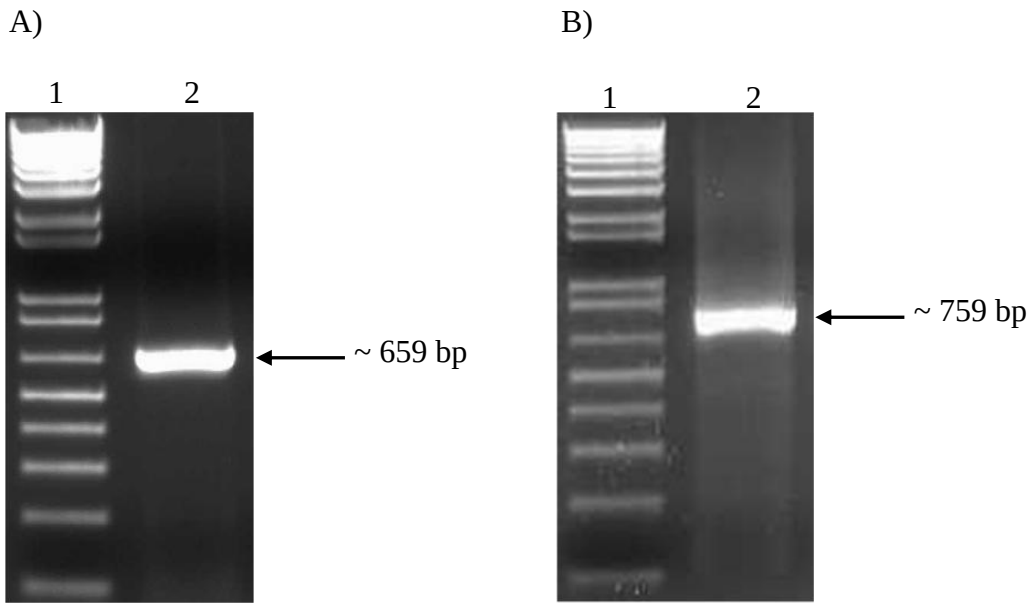


Figure 3.3 RACE-PCR products of 5’ (A) and 3’ (B) end DNA fragments of *TaSUT2* from cv. AC Andrew. Lane 1 (A, B) corresponds to DNA marker (1Kb+) used to determine the size of PCR products.

Cloning these fragments into the pGEM-T Easy vector followed by transformation of each construct with *E.coli* (DH5α) cells produced colonies. PCR screening of the colonies with specific primers designed from the 5’ and 3’ ends of each DNA fragment identified positive colonies that contained the inserted 5’ and 3’ end fragments of *TaSUT2* (Figure 3.4). Isolation of plasmids from the positive colonies and subsequent nucleotide sequencing of the inserted DNA fragments revealed that the 5’ and 3’ end fragments are 659 bp (Figure 3.5) and 759 bp (Figure 3.6) in length, respectively.

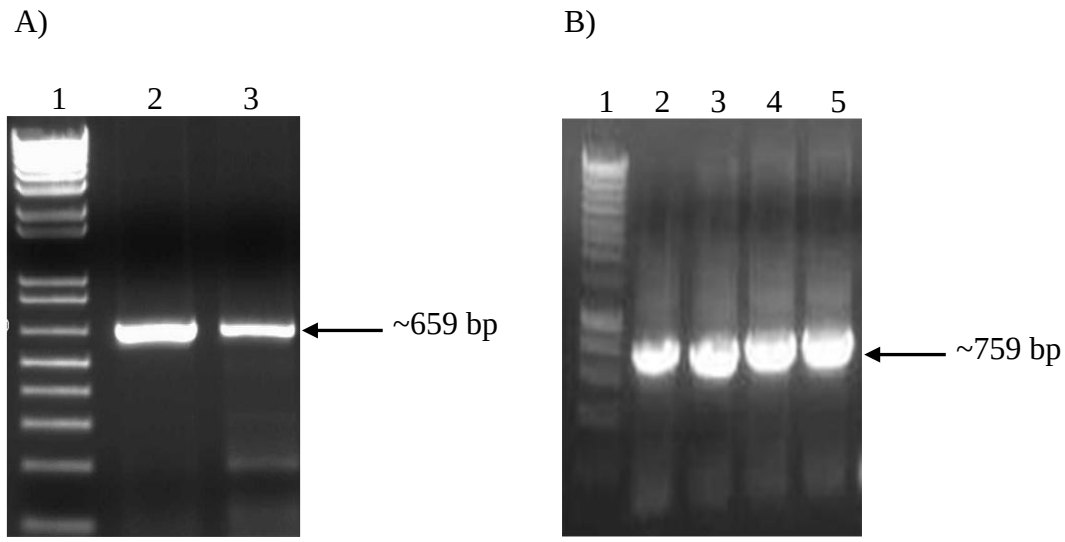


Figure 3.4 PCR screening of *E.coli* (DH5α) cells transformed with 5' (A) and 3' (B) end DNA fragments of *TaSUT2*-pGEM-T construct. Lanes 2 and 3 (A), and 2, 3, 4 and 5 (B) show DNA fragments of expected sizes of 659 bp and 759 bp, respectively, amplified from the positive colonies. Lane 1 (A, B) is for DNA marker (1Kb+).

Blasting the resulting nucleotide sequence of the putative 5' end fragment of *TaSUT2* against GenBank database showed 98% similarity with *SUT2* like sequence from wheat cv. Bobwhite (Table 3.4, Appendix 5). It also showed 90% and 81% identity with nucleotide sequences of *SUT2* from barley and rice, respectively. The nucleotide sequence of putative 3' end fragment of *TaSUT2* also exhibited 98% identity with *SUT2* like sequence from wheat cv. Bobwhite (Table 3.4, Appendix 6), and 90% and 81% identity with nucleotide sequences of *SUT2* from barley and rice, respectively.

AAGCAGTGGTATCAACGCAGAGTACGCGGGGGCACCTCCTCGGCCGCGCCGCCGCGCAA
GGTCCCGCTCCGCTCGCTGCTCCGGGCGGCGTCCGTCGCCTGCGGGGTCCAATTCGGGT
GGGCGCTGCAGCTCTCCCTGCTCACCCCGTACGTCCAAGAACTCGGCATCCCGCACGCC
TTCGCCAGCCTGGTCTGGCTCTGCGGCCCGCTGTCGGGCCTCCTCGTGCAGCCCCTGGT
GGGCCACCTCTCGGACCGCATCGCGCCGGCCAACTCCCCGCTGGGCCGCCGCCGGCCCT
TCATCGCCGCGGGGGCCGCCTCCATCGCCTTCTCCGTGCTCACCGTCGGCTTCTCCGCC
GACCTCGGCCGCCTCTTCGGGGACAACATCGTGCCCGGGTCCACCCGGTTCGGCGCCAT
CATCGTCTACCTCATCGGCTTCTGGCTGCTGGATGTCGGCAACAACGCCACCCAGGGGC
CATGCCGCGCCTTCCTCGCAGACCTCACAGAGAATGACCCGAGGAGGACCCGGATTGCC
AACGCCTACTTCTCGCTCTTCATGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGC
ATACAACGGCTGGTACAAGATATTCCCCTTCACTATTACGGAGTCCTGCTCTGTCAGCT
GCGCCAACCT

Figure 3.5 Nucleotide sequence of the 5’ end DNA fragment (659 bp) of *TaSUT2* amplified from cv. AC Andrew. The forward and reverse primers used for amplifying the fragment are underlined.

CGCCGCTTCCCTCATTGTGTTCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTA
TACCATATGCGATGGCTGCAAGTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCA
ATGGGCATTCTTAATTTATCTATTGTCATACCACAGATCATCGTGTCGCTGGGCAGCGG
GCCATGGGACCAGCTCTTCGGCGGAGGGAACGCGCCATCCTTCTGGGTGGCTGCCGCAG
CCTCCTTTGTGGGCGGGCTGGTAGCCATCCTGGGGCTCCCGCGGGCCCGGCTCGGGCCG
AAGAAGAAAACCACCCAACGATGATGATTCGAGCGAGTATACAGAGACATTGTTATGTT
GCAGACAAGGAAGGAAGGCAGGTTGTACACTGGTTCATTCTTTGCACACTGGTTCCAAG
ACACCGGGAGCGCTGCCTACACTACACTAGGCATGGGGGGTCTGTATATTGTCATCATC
AAGTTGGATTGGGTTCAAATTCAAGGCCAAAAGTAGGAGGGCTCTGGAATTTTTCTTCC
TCTTGTTACTTTCCTCCATGGGCCCTATCATTGCCTTTTTTTCGCTTGTCGGTGTGAAA
CATGGTCAGTCATATGTTGTTGTATATTCGGATATTTTGTCATTTTGTCGGAGATTATC
AATCAATCAATAAAAGTTGCAGGTTTTTGTTTGAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAGTACTCTGCGTTGATACCACTGCTTGCCCTATAGTGAGTCGTATTAG

Figure 3.6 Nucleotide sequence of the 3’ end DNA fragment (759 bp) of *TaSUT2* amplified from cv. AC Andrew. The forward and reverse primers used for amplifying the fragment are underlined.

Table 3.4 Nucleotide sequence identity between the 5' and 3' end DNA fragments of *TaSUT2* and other *SUT* genes.

Accession #	Description	Max Score[a]	E-Value[b]	Max Identtity[c]
Similarity between the 5' end fragments of *TaSUT2* and other *SUT* DNA sequences				
EU693244.1	*Triticum aestivum* cv. Bobwhite *SUT2* like	1251	0.0	94%
AJ272308.1	*Hordeum vulgare SUT2*	902	0.0	91%
GU045300.1	*Sorghum bicolor SUT4*	608	2e-170	84%
DQ072592.1	*Oryza sativa* ssp. *Indica* cv. Minghui 86 *SUT2M*	584	3e-163	83%
GQ485583.1	*Saccharum hybrid* cv. ROC22 *SUT4*	582	1e-147	85%
Similarity between the 3' end fragments of *TaSUT2* and other *SUT* DNA sequences				
EU693244.1	*Triticum aestivum* cv. Bobwhite *SUT2* like	893	0.0	98%
AJ2722308.1	*Hordeum vulgare SUT2*	854	0.0	90%
AB091672.1	*Oryza sativa SUT2*	231	4e-57	81%

[a]Maximum score indicates the amount of sequence homology between *TaSUT2* and *SUTs* from other species.
[b]E-Value determines the significance of the alignment score.
[C]Maximum identity refers to percentage of identical nucleotides between *TaSUT2* and the respective SUTs from other species.

3.3.1.3 End to end PCR-I

Assembling the partial, and 5' and 3' end DNA fragments of *TaSUT2* produced a putative full length cDNA sequence of *TaSUT2*. End to end PCR with specific forward and reverse primers designed from its 5' and 3' end produced a DNA fragment with approximate size of 1854 bp (Figure 3.7). Cloning the identified fragment into pGEM-T Easy vector and transforming *TaSUT2*-pGEM-T construct with *E.coli* (DH5α) cells

produced colonies. PCR screening of the colonies with gene specific forward and reverse primers identified positive colonies that contained the inserted putative full length cDNA of *TaSUT2*. Nucleotide sequencing of the inserted DNA fragments following isolation of plasmids from the positive colonies revealed that the putative full length *TaSUT2* is 1854 bp in length (Figure 3.8).

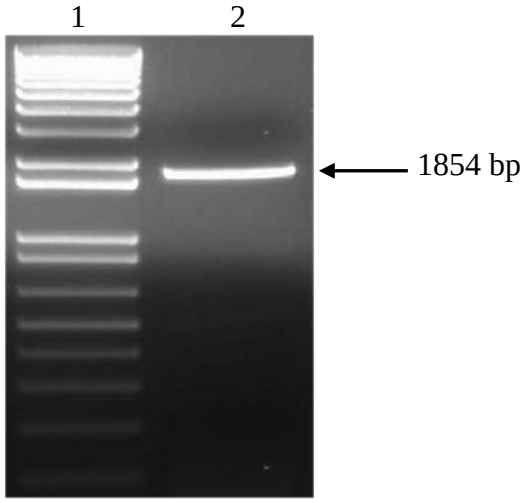


Figure 3.7 PCR product of the putative full length cDNA of *TaSUT2* amplified from cv. AC Andrew (lane 2). Lane 1 corresponds to DNA marker (1Kb+) used to determine the PCR product size.

ACGCGGGGGCACCTCCTCGGCCGCGCCGCCGCGCAAGGTCCCGCTCCGCTCGCTGCTCC
GGGCGGCGTCCGTCGCCTGCGGGGTCCAATTCGGGTGGGCGCTGCAGCTCTCCCTGCTC
ACCCCGTACGTCCAAGAACTCGGCATCCCGCACGCCTTCGCCAGCCTGGTCTGGCTCTG
CGGCCCGCTGTCGGGCCTCCTCGTGCAGCCCCTGGTGGGCCACCTCTCGGACCGCATCG
CGCCGGCCAACTCCCCGCTGGGCCGCCGCCGGCCCTTCATCGCCGCGGGGGCCGCCTCC
ATCGCCTTCTCCGTGCTCACCGTCGGCTTCTCCGCCGACCTCGGCCGCCTCTTCGGGGA
CAACATCGTGCCCGGGTCCACCCGGTTCGGCGCCATCATCGTCTACCTCATCGGCTTCT
GGCTGCTGGATGTCGGCAACAACGCCACCCAGGGGCCATGCCGCGCCTTCCTCGCAGAC
CTCACAGAGAATGACCCGAGGAGGACCCGGATTGCCAACGCCTACTTCTCGCTCTTC**AT**
GGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGATAT
TCCCCTTCACTATTACGGAGTCCTGCTCTGTCAGCTGCGCCAACCTCAAGTCCGCGTTC
CTGCTCGACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGTGACGGTGCAGGA
CAATCCTACTTTCGGAAGCGACGAGGCGGCACCTCCGAGCAGCCACGAGGAGGAGGCTT
TCCTCTTTGAGCTTTTCGGGTCGTTCAAGTACTTCACGATGCCTGTTTGGATGGTCTTG
ATCGTCACCTCGCTTACCTGGATCGGCTGGTTCCCTTTCATCCTCTTTGACACCGACTG
GATGGGCCGAGAGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAAAAGTATC
ATGACGGTGTGAGAATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCTCGGGATC
ACATCTATTGGAATGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTATGGGGTGT
CTCCAATATCATCATGGCACTGTGCTTCGTGGCGATGCTTATTATAACATACGTTGCGC
AGAACCTGGATTATGGACCTAGTGGAGCACCTCCAACCGGCATTGTCGCCGCTTCCCTC
ATTGTGTTCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTATACCATATGCGAT
GGCTGCAAGTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGCATTCTTA
ATTTATCTATTGTCATACCACAGATCATCGTGTCGCTGGGCAGCGGGCCATGGGACCAG
CTCTTCGGCGGAGGGAACGCGCCATCCTTCTGGGTGGCTGCCGCAGCCTCCTTTGTGGG
CGGGCTGGTAGCCATCCTGGGGCTCCCGCGGGCCCGGCTCGGGCCGAAGAAGAAAACCA
CCCAACGATGATGATTCGAGCGAGTATACAGAGACATTGTTATGTTGCAGACAAGGAAG
GAAGGCAGGTTGTACACTGGTTCATTCTTTGCACACTGGTTCCAAGACACCGGGAGCGC
TGCCTACACTACACTAGGCATGGGGGGTCTGTATATTGTCATCATCAAGTTGGATTGGG
TTCAAATTCAAGGCCAAAAGTAGGAGGGCTCTGGAATTTTTCTTCCTCTTGTTACTTTC
CTCCATGGGCCCTATCATTGCCTTTTTTTCGCTTGTCGGTGTGAAACATGGTCAGTCAT
ATGTTGTTGTATATTCGGATATTTTGTCATTTTGTCGGAGATTATCAATCAATCAATAA
AAGTTGCAGGTTTTTGTTTGAAAAA

Figure 3.8 Nucleotide sequence of the putative full length cDNA of *TaSUT2* (1854 bp). The forward and reverse primers used for amplifying the full length cDNA fragment are underlined. The predicted coding sequence of *TaSUT2* (957 bp) is shown in bold.

Blasting the nucleotide sequence of the putative full length cDNA of *TaSUT2* against GenBank database showed 99% identity with the *SUT2* like sequence from wheat, cv. Bobwhite (Table 3.5), and also 92% and 81% identity with nucleotide sequences of *SUT2* genes from barley and rice, respectively (Table 3.5).

Table 3.5 Nucleotide sequence identity between full sequence of *TaSUT2* and *SUTs* from other species.

Accession #	Description	Max Score[a]	E-Value[b]	Max Identity[c]
EU693244.1	*Triticum aestivum* cv. Bobwhite *(SUT2* like*)*	2983	0.0	99%
AJ272308.1	*Hordeum vulgare SUT2*	2518	0.0	92%
GU045300.1	*Sorghum bicolor SUT4*	1086	0.0	81%
AK109461.1	*Oryza sativa* ssp. *Japonica SUT2*	1077	0.0	81%
DQ072592.1	*Oryza sativa* ssp. *Indica* cv. Minghui 86 *SUT2M*	1066	0.0	80%
AY639018.1	*Zea mays SUT2*	1009	0.0	80%

[a]Maximum score indicates the amount of sequence homology between *TaSUT2* and *SUTs* from other species.
[b]E-Value determines the significance of the alignment score.
[C]Maximum identity refers to percentage of identical nucleotides between *TaSUT2* and the respective SUTs from other species.

Analysis of the putative full length fragment of *TaSUT2* with open reading frame (ORF) finder software (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) revealed that the DNA fragment has a coding sequence of 957 bp in length (Figure 3.8). Translating the putative coding sequence into amino acid sequence using a translator tool from JustBio (http://www.justbio.com/index) produced a protein with 319 amino acids (Figure 3.9). Blasting the resulting amino acid sequence against GenBank database confirmed that the identified putative TaSUT2 is a member of the GPH cation symporter family to which all other known plant sucrose-proton symporters belong (Figure 3.10, Table 3.6). However,

this putative TaSUT2 protein lacked a complete protein domain structure that is characteristic to all members of the GPH cation symporter family (Figure 3.10).

MALGNILGYATGAYNGWYKIFPFTITESCSVSCANLKSAFLLDIIILAITTYVSVVTVQ
DNPTFGSDEAAPPSSHEEEAFLFELFGSFKYFTMPVWMVLIVTSLTWIGWFPFILFDTD
WMGREIYRGSPEIVADTQKYHDGVRMGSFGLMLNSVVLGITSIGMEKLCRKWGAGLVWG
VSNIIMALCFVAMLIITYVAQNLDYGPSGAPPTGIVAASLIVFTILGAPLSVTYSIPYA
MAASRVENLGLGQGLAMGILNLSIVIPQIIVSLGSGPWDQLFGGGNAPSFWVAAAASFV
GGLVAILGLPRARLGPKKKTTQR*

Figure 3.9 Amino acid sequence of the putative TaSUT2 resulted from the translation of the putative coding nucleotide sequence of *TaSUT2*.

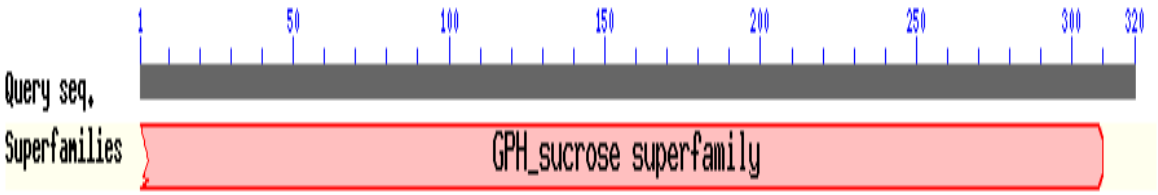


Figure 3.10 Protein domain structure of the initialy identified TaSUT2 as compared to other members of the GPH cation symporter family. The initially identified TaSUT2 contains incomplete protein domains (see left end).

Table 3.6 Amino acid sequence similarity between the putative TaSUT2 and SUT proteins from other species.

Accession #	Description	Max Score[a]	E -Value[b]
CAB7588.1	*Hordeum vulgare* SUT2	625	0.0
ACV95498.1	*Saccharum hybrid* cv. ROC22 SUT4	518	0.0
ADY11193.1	*Oryza sativa* ssp. *Japonica* SUT2	516	0.0
AAT51689.1	*Zea mays* SUT2	494	1e-172
CAG70682.1	*Datisca glomerata* putative SUT4	421	1e-143
BAI60050.1	*Nicotiana tabacum* SUT4	419	5e-143

[a]Maximum score indicates the amount of sequence homology between *TaSUT2* and *SUTs* from other species.
[b]E-Value determines the significance of the alignment score.

3.3.1.4 Extension of the 5' end fragment

Amplification of cDNA samples using a specific primer designed from the incomplete 5' end fragment (Appendix 7) and universal inner and outer primers (Clontech) produced a DNA fragment with approximate size of ~261 bp (Figure 3.11A). Cloning the ~261 bp DNA fragment into the pGEM-T Easy vector and transformation of *E.coli* (DH5α) cells with the resulting construct produced colonies. Screening of the colonies with PCR identified the positive ones that contained the extended 5' end DNA fragment (Figure 3.11B). Plasmid isolation from the positive colonies and subsequent nucleotide sequencing of the inserted DNA fragment revealed that the extended 5' end fragment is 261 bp in length (Figure 3.12).

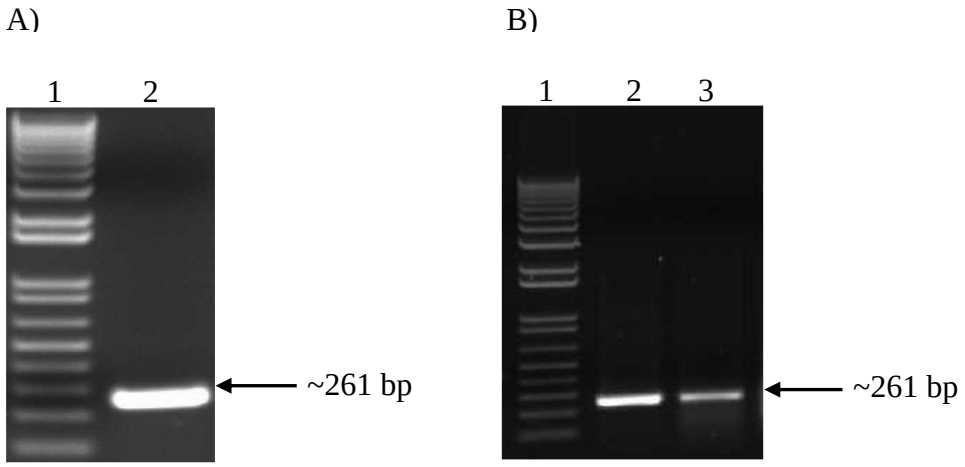


Figure 3.11 PCR product of the DNA fragment extended from the 5' end of the putative full length cDNA of *TaSUT2* isolated from cv. AC Andrew (A; lane 2). PCR screening of *E.coli* (DH5α) cells transformed with the extended 5' end DNA fragment of *TaSUT2*-pGEM-T construct (B); lanes 2 and 3 show DNA fragments of expected size (261 bp) amplified from the positive colonies. Lane 1 (A, B) corresponds to DNA marke (1Kb+).

AAGCAGTGGTATCAACGCAGAGTACGCGGGGAGAGCGATAGATACTTTGCTTTCCCCTC
CCTCCATCCCTTCCCTTCCCTCTCTCCCAGGTCAGGCCAGGTCTCCCACAGAATTCAGC
CAGGAGGCGTGGGGGGAGCGGGCAGCCAAGAAAGTCCAACACCACTCATCAAGGGCGAT
GCCGCCGCGGCGGCCCAACACCGGCGCTGGGGGCGGGGGCACCTCCTCGGCCGCGCCGC
CGCGCAAGGTCCCGCTCCGCTCGCT

Figure 3.12 Nucleotide sequence of the DNA fragment extended from the 5' end of putative cDNA of *TaSUT2*. The forward and reverse primers used to amplify the fragment are underlined.

3.3.1.5 End to end PCR II

Assembling the newly identified 5' end fragment (containing the extension) with the previously identified 3' and partial fragments of *TaSUT2* produced a putative complete full length sequence of *TaSUT2* with approximate size of 2036 bp in length. End to end PCR of the cDNA using the gene specific forward and reverse primers produced DNA

fragment of expected size (2036 bp; Figure 3.13A, Appendix 8). Cloning the putative complete full length cDNA sequence of *TaSUT2* into the pGEM-T Easy vector followed by transformation of *E.coli* (DH5α) cells with the resulting construct produced colonies. PCR screening of the colonies with primers used for end to end PCR identified positive colonies that contained the putative complete full length cDNA of *TaSUT2* (Figure 3.13B). Sequencing of the putative complete full length *TaSUT2* cDNA fragment contained in the plasmid isolated from the positive colonies confirmed that the fragment has a size of 2036 bp in length (Figure 3.14). Search for identical nucleotide sequences by blasting the putative *TaSUT2* cDNA sequence against GenBank database revealed that the fragment has 97% identity with the nucleotide sequence of *SUT2* like gene from wheat cv. Bobwhite (Table 3.7), 92% identity with barley and 80% with rice and maize *SUT2*s, respectively (Table 3.7).

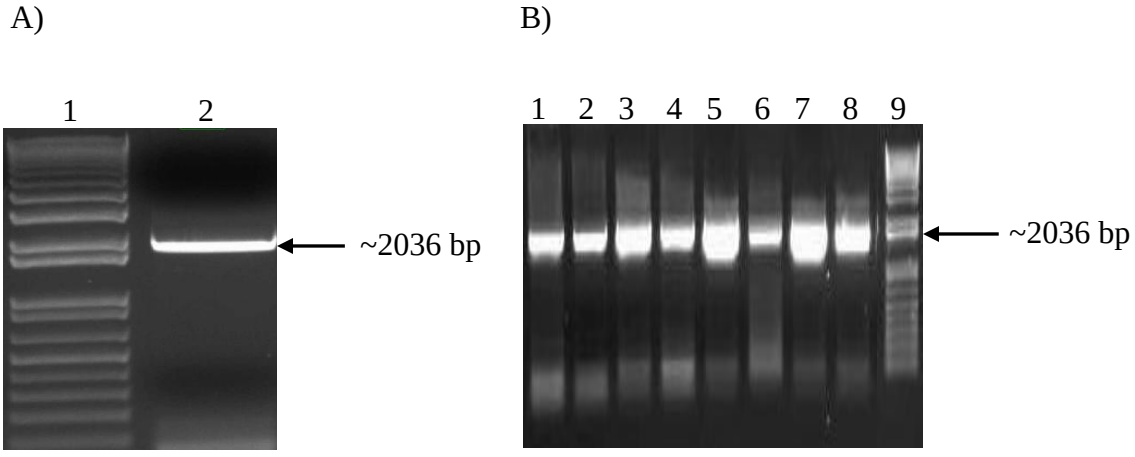


Figure 3.13 End to end PCR product of the putative complete full length cDNA of *TaSUT2* amplified from cv. AC Andrew (A; lane 2). PCR screening of *E.coli* (DH5α) cells transformed with the putative complete cDNA fragment of *TaSUT2*-pGEM-T construct; lanes 1, 2, 3, 4, 5, 6, 7, and 8 show DNA fragments of expected size (2036 bp) amplified from the positive colonies (B). Lanes 1 (A) and 9 (B) corresponds to DNA marker (1Kb+).

ACGCGGGGAGAGCGATAGATACTTTGCTTTCCCCTCCCTCCATCCCTTCCCTTCCCTCT
CTCCCAGGTCAGGCCAGGTCTCCCACAGAATTCAGCCAGGAGGCGTGGGGGGAGCGGGC
AGCCAAGAAAGTCCAACACCACTCATCAAGGGCG**ATGCCGCCGCGGCGGCCCAACACCG**
GCGCCGGGGGCGGGGGCACCTCCTCGGCCGCGCCGCCGCGCAAGGTCCCGCTCCGCTCG
CTGCTCCGGGCGGCGTCCGTCGCCTGCGGGGTCCAATTCGGGTGGGCGCTGCAGCTCTC
CCTGCTCACCCCGTACGTCCAAGAACTCGGCATCCCGCACGCCTTCGCCAGCCTGGTCT
GGCTCTGCGGCCCGCTGTCGGGCCTCCTCGTGCAGCCCCTGGTGGGCCACCTCTCGGAC
CGCATCGCGCCGGCCAACTCCCCGCTGGGCCGCCGCCGGCCCTTCATCGCCGCGGGGGC
CGCCTCCATCGCCTTCTCCGTGCTCACCGTCGGCTTCTCCGCCGACCTCGGCCGCCTCT
TCGGGGACAACATCGTGCCCGGGTCCACCCGGTTCGGCGCCATCATCGTCTACCTCATC
GGCTTCTGGCTGCTGGATGTCGGCAACAACGCCACCCAGGGGCCATGCCGCGCCTTCCT
CGCAGACCTCACAGAGAATGACCCGAGGAGGACCCGGATTGCCAACGCCTACTTCTCGC
TCTTCATGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTAC
AAGATATTCCCCTTCACTATTACGGAGTCCTGCTCTGTCAGCTGCGCCAACCTCAAGTC
CGCGTTCCTGCTCGACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGTGACGG
TGCAGGACAATCCTACTTTCGGAAGCGACGAGGCGGCACCTCCGAGCAGCCACGAGGAG
GAGGCTTTCCTCTTTGAGCTTTTCGGGTCGTTCAAGTACTTCACGATGCCTGTTTGGAT
GGTCTTGATCGTCACCTCGCTTACCTGGATCGGCTGGTTCCCTTTCATCCTCTTTGACA
CCGACTGGATGGGCCGAGAGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAA
AAGTATCATGACGGTGTGAGAATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCT
CGGGATCACATCTATTGGAATGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTAT
GGGGTGTCTCCAATATCATCATGGCACTGTGCTTCGTGGCGATGCTTATTATAACATAC
GTTGCGCAGAACCTGGATTATGGACCTAGTGGAGCACCTCCAACCGGCATTGTCGCCGC
TTCCCTCATTGTGTTCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTATACCAT
ATGCGATGGCTGCAAGTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGC
ATTCTTAATTTATCTATTGTCATACCACAGATCATCGTGTCGCTGGGCAGCGGGCCATG
GGACCAGCTCTTCGGCGGAGGGAACGCGCCATCCTTCTGGGTGGCTGCCGCAGCCTCCT
TTGTGGGCGGGCTGGTAGCCATCCTGGGGCTCCCGCGGGCCCGGCTCGGGCCGAAGAAG
AAAACCACCCAACGATGATGATTCGAGCGAGTATACAGAGACATTGTTATGTTGCAGAC
AAGGAAGGAAGGCAGGTTGTACACTGGTTCATTCTTTGCACACTGGTTCCAAGACACCG
GGAGCGCTGCCTACACTACACTAGGCATGGGGGGTCTGTATATTGTCATCATCAAGTTG
GATTGGGTTCAAATTCAAGGCCAAAAGTAGGAGGGCTCTGGAATTTTTCTTCCTCTTGT
TACTTTCCTCCATGGGCCCTATCATTGCCTTTTTTTCGCTTGTCGGTGTGAAACATGGT
CAGTCATATGTTGTTGTATATTCGGATATTTTGTCATTTTGTCGGAGATTATCAATCAA
TCAATAAAAGTTGCAGGTTTTTGTTTGAAA

Figure 3.14 Nucleotide sequence of the complete full length cDNA of *TaSUT2* (2036 bp). The primers used to amplify the full length cDNA fragment are underlined, and the ORF is shown in bold.

Table 3.7 Nucleotide sequence identity between the full length *TaSUT2* and *SUTs* from other species.

Accession #	Description	Max Score[a]	E-Value[b]	Max Identity[c]
EU693244.1	*Triticum aestivum* cv. Bobwhite *(SUT2)*	3099	0.0	97%
AJ272308.1	*Hordeum vulgare SUT2*	2588	0.0	92%
GU045300.1	*Sorghum bicolor SUT4*	1086	0.0	81%
DQ072592.1	*Oryza sativa* ssp. *Indica* cv. Minghui 86 *SUT2M*	1075	0.0	80%
AB091672.1	*Oryza sativa* ssp. *Japonica SUT2*	1059	0.0	80%
AY639018.1	*Zea mays SUT*2	1009	0.0	80%

[a]Maximum score indicates the amount of sequence homology between *TaSUT2* and *SUTs* from other species.
[b]E-Value determines the significance of the alignment score.
[C]Maximum identity refers to percentage of identical nucleotides between *TaSUT2* and the respective SUTs from other species.

Analysis of the newly identified putative complete full length cDNA of *TaSUT2* with the ORF finder software revealed that the DNA fragment has a coding nucleotide sequence of 1518 bp (Figure 3.14). Translating the resulting coding sequence into amino acid sequence using a translator tool from JustBio produced a protein with 506 amino acids (Figure 3.15). Blasting the resulting amino acid sequence against GenBank database confirmed that the identified putative TaSUT2 is a new member of the GPH cation symporter family to which all other plant sucrose-proton symporters belong, and it contained all the required protein domains (Figure 3.16). Protein topology prediction using TMpred program (http://www.ch.embnet.org/software/TMPRED_form.html; Hofmann and Stoffel, 1993) indicated that TaSUT2 contains 12 transmembrane helices (Figure 3.15), which is coined as a distinct characteristic feature of all known SUTs (Marger and Saier, 1993; Shiratake, 2007). Furthermore, TaSUT2 contains the consensus

sequence derived from the highly conserved region of functional plant SUTs (Aoki et al., 2003; Figure 3.17). The histidine residue, considered to be responsible for sucrose binding in all known SUTs, is also present in TaSUT2 (Figure 3.17). Protein BLAST showed that TaSUT2 has 97%, 85% and 83% similarity with SUT2 proteins from barley (*Hordeum vulgare*; HvSUT2), rice (*Oryza sativa*; OsSUT2) and maize (*Zea mays*; ZmSUT2), respectively (Table 3.8).

Table 3.8 Amino acid sequence similarity between TaSUT2 and SUT proteins from other species.

Accession #	Description	Max Score[a]	E-Value[b]	Similarity (%)
CAB7588.1	*Hordeum vulgare* SUT2	949	0.0	97%
XP 003577278.1	*Brachypodium distachyon* SUT2	845	0.0	89%
ACX71839.1	*Sorghum bicolor* SUT4	813	0.0	83%
ADY11193.1	*Oryza sativa* ssp. *Japonica* SUT2	516	0.0	85%
AAT51689.1	*Zea mays* SUT2	796	0.0	83%
CAG70682.1	*Datisca glomerata* putative SUT4	651	0.0	65%
BAI60050.1	*Nicotiana tabacum* SUT4	649	0.0	65%

[a]Maximum score indicates the amount of sequence homology between *TaSUT2* and *SUTs* from other species.
[b]E-Value determines the significance of the alignment score.

```
                                                        I
MPPRRPNTGAGGGGTSSAAPPRKVPLRSLLRAASVACGVQFGWALQLSLLTPYVQELGI
           II                                                III
PHAFASLVWLCGPLSGLLVQPLVGHLSDRIAPANSPLGRRRPFIAAGAASIAFSVLTVG
                               IV
FSADLGRLFGDNIVPGSTRFGAIIVYLIGFWLLDVGNNATQGPCRAFLADLTENDPRRT
          V                                                  VI
RIANAYFSLFMALGNILGYATGAYNGWYKIFPFTITESCSVSCANLKSAFLLDIIILAI
                                                             VII
TTYVSVVTVQDNPTFGSDEAAPPSSHEEEAFLFELFGSFKYFTMPVWMVLIVTSLTWIG
                                                    VIII
WFPFILFDTDWMGREIYRGSPEIVADTQKYHDGVRMGSFGLMLNSVVLGITSIGMEKLC
                       IX                                X
RKWGAGLVWGVSNIIMALCFVAMLIITYVAQNLDYGPSGAPPTGIVAASLIVFTILGAP
                                  XI
LSVTYSIPYAMAASRVENLGLGQGLAMGILNLSIVIPQIIVSLGSGPWDQLFGGGNAPS
       XII
FWVAAAASFVGGLVAILGLPRARLGPKKKTTQR*
```

Figure 3.15 Amino acid sequence of the putative TaSUT2. The predicted 12 transmembrane helices are underlined.

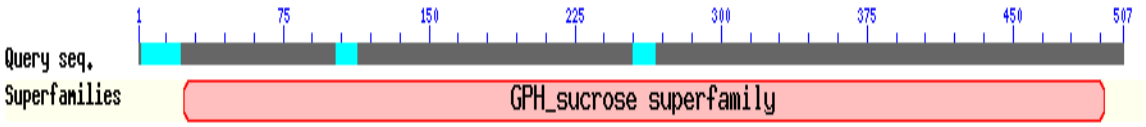


Figure 3.16 Protein domain structure of the putative TaSUT2 as compared to other members of the GPH cation symporter family. The putative TaSUT2 contains all the required protein domains.

```
Consensus                                                  A.G1Q2GWALQLSLLTPY1Q.LG1
TaSUT2          MPPRRRPNTGAGGGGTSSAAP-PRKVPLRSLLRAASVACGVQFGWALQLSLLTPYVQELGI
HvSUT2          MPPRRPNTGGGGGTSSSAAPAPRKVPLRSLLRAASVACGVQFGWALQLSLLTPYVQELGI
OsSUT2          -MPRRPS--GGGGGAGPAAAAVRKVPLRKLLRAASVACGVQFGWAPQLSLLTPYVQELGI
ZmSUT2          ------M--PPRTAPAATSTPPRKVPLRKLLRAASVACGVQFGWALQLSLLTPYVQELGI
                              ..   ****** ********************** **************

Consensus       .H......WLCGP1.G..VQP..G..SD.---.....GRR..2I..G
TaSUT2          PHAFASLVWLCGPLSGLLVQPLVGHLSDRIAPANSPLGRRRPFIAAGAASIAFSVLTVGF
HvSUT2          PHAFASLVWLCGPLSGLLVQPLVGHLSDRITPANSPLGRRRPFIAAGAASIAFAVLTVGF
OsSUT2          PHAFASLVWLCGPLSGLLVQPLVGHLSDRIAPAASPLGRRRPFIAAGAASIAAAVLTVRF
ZmSUT2          PHAFASLVWLCGPLSGLLVQPLVGHLSDRIGPAASPLGRRRPFIAAGAACIAAAVLTVGF
                **************************** ** ***************.** .**** *

TaSUT2          SADLGRLFGDNIVPGSTRFGAIIVYLIGFWLLDVGNNATQGPCRAFLADLTENDPRRTRI
HvSUT2          SADLGRLFGDNVVPGSTRIGAIIVYLVGFWLLDVGNNATQGPCRAFLADLTENDPRRTRI
OsSUT2          SADLGRIFGDSITPGSTRLGAITAYLVGFWLLDVGNNATQGPCRAFPADLTENDPKRTRI
ZmSUT2          SADLGRLFGDDVTPGSTRLGAICVYLVGFWLLDVGNNGTQGPCRAFLADLTENDPRRTRI
                ******.*** . ***** ***  **.********** ******** ********.****

TaSUT2          ANAYFSLFMALGNILGYATGAYNGWYKIFPFTITESCSVSCANLKSAFLLDIIILAITTY
HvSUT2          ANAYFSLFMALGNILGYATGAYNGWYKIFPFTITGSCGVSCANLNSAFLLDIIILAITTY
OsSUT2          ANAYFSLFMALGNILGYATGAYSGWYKIFPFTVTPSCSISCANLKSAFLLDIIILVVTTC
ZmSUT2          ANAYFSLFMALGNILGYATGAYSGWYSIFPFTVTESCGISCANLKSAFLLDIIVLVITTY
                ********************** *** *****.* ** .***** ********.* .**

TaSUT2          VSVVTVQDNPTFGSDEAAPPSSHEEEAFLFELFGSFKYFTMPVWMVLIVTSLTWIGWFPF
HvSUT2          ISVATVQDNPTFGSDEAAPPSSHEEEAFLFELFGSFKYFTMPVWMVLIVTSLTWVGWFLF
OsSUT2          ITVASVQEPQSFGSDEADHPSTEQ-EAFLWELFGSFRYFTLPVWMVLIVTALTWIGWFPF
ZmSUT2          TTVTSVQEPQTFGSDEAQNPGAEQ-EAFLWELFGSLRYFTLPIWMVLIVTALTWMAWFPF
                 .* .**.  .******  * . . **** ***** .***.*.*******.***. ** *

TaSUT2          ILFDTDWMGREIYRGSPEIVADTQKYHDGVRMGSFGLMLNSVVLGITSIGMEKLCRKWGA
HvSUT2          ILFDTDWMGREIYRGSPEIVADTQKYHDGVRMGSFGLMLNSVVLGITSIGMEKLCRKWGA
OsSUT2          ILFDTDWMGREIYRGSPDDPSITQSYHDGVRMGSFGLMLNSVLLGFTSIVLEKLCRKWGA
ZmSUT2          TLFDTDWMGREIYRGSPDNPGETQRYHDGVRMGSFGLMLNSVVLGFTSVVLEKLCRKWGA
                 ****************.    ** *****************.** **. .*********

TaSUT2          GLVWGVSNIIMALCFVAMLIITYVAQNLDYGPSGAPPTGIVAASLIVFTILGAPLSVTYS
HvSUT2          GLVWGVSNIIMALCFVAMLIITYVAQNLDYGPSGAPPTGIVAASLIVFTILGAPLSITYS
OsSUT2          GLVWGVSNILMALCFVAMLVITYVAKNMDYPPSGVPPTGIVIASLVVFTILGAPLAITYS
ZmSUT2          GLVWGVSNILMTLCFLAMLVITYVAKNMDYPSSGAPPTGIVVASLVVFTILGAPLAITYS
                *********.*.***.***.*****.*.**  ** ****** ***.*********..***

TaSUT2          IPYAMAASRVENLGLGQGLAMGILNLSIVIPQIIVSLGSGPWDQLFGGGNAPSFWVAAAA
HvSUT2          IPYAMAASRVENLGLGQGLAMGILNLSIVIPQIIVSLGSGPWDQLFGGGNAPSFWVAAAA
OsSUT2          IPYAMAASRVENLGLGQGLAMGILNLAIVIPQVIVSLGSGPWDQLFGGGNAPAFAVAAAA
ZmSUT2          IPYAMAASRVENLGLPGQGLAMGILNLAIVIPQVIVSLGSGPWDQLFGGGNAPAFAVAAGA
                ************** ***********.*****.******************.* *** *

TaSUT2          SFVGGLVAILGLPRARLGPKKKTTQRX---
HvSUT2          SFVGGLVAILGLPRARLGPKKKTTQR----
OsSUT2          SFIGGLVAILGLPRARIA---SRRRGHR--
ZmSUT2          SFIGGLVAILGLPRARIASSSSRRRGGTHR
                **.**************.       .
```

Figure 3.17 Alignment of the amino acid sequences of TaSUT2 with other SUT2 proteins derived from other cereal crops: barley (*Hordeum vulgare*; HvSUT2), rice (*Oryza sativa*; OsSUT2) and maize (*Zea mays*; ZmSUT2). Similarity in nucleotides across all the sequences is indicated by stars. The sequence of TaSUT2 is compared with the CONSENSUS sequence derived from highly conserved region of functional plant SUTs where 1 = I, L or V; 2 = F, W or Y. The histidine residue, which is important for the functionality of SUT proteins, is shown in bold and underlined.

3.3.2 Identification of *TaSUT2* in the progenitors of A, B and D genomes

The presence of *TaSUT2* in A, B and D genomes was examined by using the three diploid progenitors of hexaploid wheat, *Triticum urartu* (A genome donor), *Aegilops speltoides* (B genome donor) and *Aegilops tauschi* (D genome donor). Amplification of the cDNA samples prepared from the young leaf tissues of the three diploid progenitors of wheat using forward and reverse primers designed for amplifying the full length of *TaSUT2* produced DNA fragments of expected size of approximately 2036 bp (Figure 3.18). Cloning of the DNA fragments amplified from the diploid progenitors into pGEM-T Easy vector and subsequent transformation of *E.coli* (DH5α) cells with the respective constructs produced colonies. PCR screening of the colonies using the specific primers described above identified positive colonies that contained the respective *TaSUT2* fragments. Plasmid isolation from the positive colonies and subsequent nucleotide sequencing revealed that the putative full length *SUT2* DNA fragments from *Triticum urartu*, *Aegilops speltoides* and *Aegilops tauschii* are 2036, 2018 and 2042 bp in length, respectively (Figure 3.19). Analysis of the resulting nucleotide sequences with ORF finder indicated that *TuSUT2A* and *AesSUT2* each has a coding sequence of 1518 bp in length, whereas that of *AetSUT2* is 1524 bp (Figure 3.19). The amino acid sequences of the SUT2s derived from the three genome progenitors showed over 99% similarity to one another (Figure 3.20). The full length sequence of the *TaSUT2* initially identified from cv. AC Andrew by RACE-PCR (see section 3.3.1.5) showed 100% identity with that obtained from *Triticum urartu*, implying that it originated from the A genome of cv. AC Andrew.

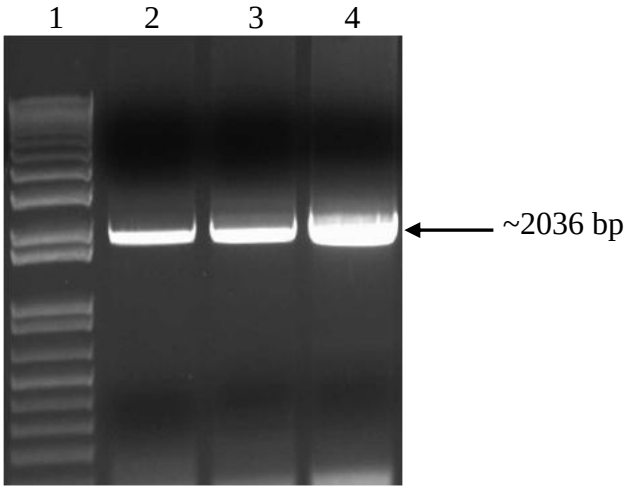


Figure 3.18 PCR products of the full length DNA fragments of *SUT2* (~ 2036 bp) amplified from the three diploid progenitors of hexaploid wheat; *Triticum urartu* (A genome donor; lane 2), *Aegilops speltoides* (B genome donor; lane 3) and *Aegilops tauschii* (D genome donor; lane 4). Lane 1 represents the DNA marker (1Kb+).

```
TaSUT2A         ACGCGGGGAGAGCGATAGATACTTTGCTTTCCCCTCCCTCCATCCCTTCCCTTCCCTCTC
TaSUT2B         ACGCGGGGAGAGCGATAGATACTTTGCTTTCCCCTCCCTCCATCCCTTCCCTTCCCTCTC
TaSUT2D         ACGCGGGGAGAGCGATAGATACTTTGCTTTCCCCTCCCTCCATCCCTTCCCTTCCCTCTC
                ************************************************************

TaSUT2A         TCCCAGGTCAGGCCAGGTCTCCCACAGAATTCAGCCAGGAGGCGTGGGGGGAGCGGGCAG
TaSUT2B         TCCCAGGTCAGGCCAGGTCTCCCACAGAATTCAGCCAGGAGGCGTGGGGGGAGCGGGCAG
TaSUT2D         TCCCAGGTCAGGCCAGGTCTCCCGCAGAATTCAGCCAGGAGGCGTGGGGGGAGCGGGCAG
                ********************** ***********************************

TaSUT2A         CCAAGAAAGTCCAACACCACTCATCAAGGGCGATGCCGCCGCGGCGGCCCAACACCGGCG
TaSUT2B         CCAAGAAAGTCCAACACCACTCATCAAGGGCGATGCCGCCGCGGCGGCCCAACACCGGCG
TaSUT2D         CCAAGAAAGTCCAACACCGCTCATCAAGGGCGATGCCGCCGCGCCGGCCCAACACCGG--
                ****************** ********************** **************

TaSUT2A         CCGGGGGCGGGGGCACCTCCTCGGCCGCG---------CCGCCGCGCAAGGTCCCGCTCC
TaSUT2B         CCGGGGGCGGGGGCACCTCCTCGGCCGCG---------CCGCCGCGCAAGGTCCCGCTCC
TaSUT2D         -CGGGGGCGGGGGCACCTCCTCGGCCGCGCCGCCGCCTCCGCCGCGCAAGGTCCCGCTCC
                 ***************************         **********************

TaSUT2A         GCTCGCTGCTCCGGGCGGCGTCCGTCGCCTGCGGGGTCCAATTCGGGTGGGCGCTGCAGC
TaSUT2B         GCTCGCTGCTCCGGGCGGCGTCCGTCGCCTGCGGGGTCCAATTCGGGTGGGCGCTGCAGC
TaSUT2D         GCTCGCTGCTCCGGGCGGCCTCCGTCGCCTGCGGGGTCCAATTCGGGTGGGCGCTGCAGC
                ****************** ****************************************
```

Continued

```
TaSUT2A         TCTCCCTGCTCACCCCGTACGTCCAAGAACTCGGCATCCCGCACGCCTTCGCCAGCCTGG
TaSUT2B         TCTCCCTGCTCACCCCGTACGTCCAAGAACTCGGCATCCCGCACGCCTTCGCCAGCCTGG
TaSUT2D         TCTCCCTGCTCACCCCCTACGTCCAAGAGCTCGGCATCCCGCACGCCTTCGCCAGCCTGG
                ********** ********** ********** ********** ********** ***

TaSUT2A         TCTGGCTCTGCGGCCCGCTGTCGGGCCTCCTCGTGCAGCCCCTGGTGGGCCACCTCTCGG
TaSUT2B         TCTGGCTCTGCGGCCCGCTGTCGGGCCTCCTCGTGCAGCCCCTGGTGGGCCACCTCTCGG
TaSUT2D         TCTGGCTCTGCGGCCCGCTGTCGGGCCTCCTCGTGCAGCCCCTGGTGGGCCACCTCTCGG
                ********** ********** ********** ********** ********** ***

TaSUT2A         ACCGCATCGCGCCGGCCAACTCCCCGCTGGGCCGCCGCCGGCCCTTCATCGCCGCGGGGG
TaSUT2B         ACCGCATCGCGCCGGCCAACTCCCCGCTGGGCCGCCGCCGGCCCTTCATCGCCGCGGGGG
TaSUT2D         ACCGCATCGCGCCGGCCAACTCGCCGCTCGGCCGCCGCCGGCCCTTCATCGCCGCGGGGG
                ********** ********** ********** ********** ********** ***

TaSUT2A         CCGCCTCCATCGCCTTCTCCGTGCTCACCGTCGGCTTCTCCGCCGACCTCGGCCGCCTCT
TaSUT2B         CCGCCTCCATCGCCTTCTCCGTGCTCACCGTCGGCTTCTCCGCCGACCTCGGCCGCCTCT
TaSUT2D         CCGCCTCCATCGCCTTCTCCGTGCTCACCGTCGGCTTCTCCGCCGACCTGGGCCGCCTCT
                ********** ********** ********** ********** ********** ***

TaSUT2A         TCGGGGACAACATCGTGCCCGGGTCCACCCGGTTCGGCGCCATCATCGTCTACCTCATCG
TaSUT2B         TCGGGGACAACATCGTGCCCGGGTCCACCCGGTTCGGCGCCATCATCGTCTACCTCATCG
TaSUT2D         TCGGGGACAACATCGTGCCCGGGTCCACCCGGTTCGGCGCCATCATCGTCTACCTCATCG
                ********** ********** ********** ********** ********** ***

TaSUT2A         GCTTCTGGCTGCTGGATGTCGGCAACAACGCCACCCAGGGGCCATGCCGCGCCTTCCTCG
TaSUT2B         GCTTCTGGCTGCTGGATGTCGGCAACAACGCCACCCAGGGGCCATGCCGCGCCTTCCTCG
TaSUT2D         GCTTCTGGCTGCTGGATGTGGGCAACAACGCCACCCAGGGGCCATGCCGCGCCTTCCTCG
                ********** ********** ********** ********** ********** ***

TaSUT2A         CAGACCTCACAGAGAATGACCCGAGGAGGACCCGGATTGCCAACGCCTACTTCTCGCTCT
TaSUT2B         CAGACCTCACAGAGAATGACCCGAGGAGGACCCGGATTGCCAACGCCTACTTCTCGCTCT
TaSUT2D         CAGACCTCACAGAGAATGACCCGAGGAGGACCCGGATTGCCAACGCCTACTTCTCGCTCT
                ********** ********** ********** ********** ********** ***

TaSUT2A         TCATGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGA
TaSUT2B         TCATGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGA
TaSUT2D         TCATGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGA
                ********** ********** ********** ********** ********** ***

TaSUT2A         TATTCCCCTTCACTATTACGGAGTCCTGCTCTGTCAGCTGCGCCAACCTCAAGTCCGCGT
TaSUT2B         TATTCCCCTTCACTATTACGGAGTCCTGCTCTGTCAGCTGCGCCAACCTCAAGTCCGCGT
TaSUT2D         TATTCCCCTTCACTATTACCGAGTCCTGCTCTGTCAGCTGCGCCAACCTCAAGTCCGCGT
                ********** ********** ********** ********** ********** ***

TaSUT2A         TCCTGCTCGACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGTGACGGTGCAGG
TaSUT2B         TCCTGCTCGACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGTGACGGTGCAGG
TaSUT2D         TCCTGCTCGACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGTGACGGTGCAGG
                ********** ********** ********** ********** ********** ***

TaSUT2A         ACAATCCTACTTTCGGAAGCGACGAGGCGGCACCTCCGAGCAGCCACGAGGAGGAGGCTT
TaSUT2B         ACAATCCTACTTTCGGAAGCGACGAGGCGGCACCTCCGAGCAGCCACGAGGAGGAGGCTT
TaSUT2D         ACAATCCTACTTTCGGAAGCGACGAGGCGGCGCCTCCGAGCAGCCACGAGGAGGAGGCTT
                ********** ********** ********** ********** ********** ***

TaSUT2A         TCCTCTTTGAGCTTTTCGGGTCGTTCAAGTACTTCACGATGCCTGTTTGGATGGTCTTGA
TaSUT2B         TCCTCTTTGAGCTTTTCGGGTCGTTCAAGTACTTCACGATGCCTGTTTGGATGGTCTTGA
TaSUT2D         TCCTCTTTGAGCTTTTCGGGTCGTTCAAGTACTTCACGATGCCTGTTTGGATGGTCTTGA
                ********** ********** ********** ********** ********** ***
```

Continued

```
TaSUT2A         TCGTCACCTCGCTTACCTGGATCGGCTGGTTCCCTTTCATCCTCTTTGACACCGACTGGA
TaSUT2B         TCGTCACCTCGCTTACCTGGATCGGCTGGTTCCCTTTCATCCTCTTTGACACCGACTGGA
TaSUT2D         TCGTCACCTCGCTTACCTGGATCGGCTGGTTCCCTTTCATCCTCTTTGACACCGACTGGA
                ************************************************************

TaSUT2A         TGGGCCGAGAGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAAAAGTATCATG
TaSUT2B         TGGGCCGAGAGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAAAAGTATCATG
TaSUT2D         TGGGCCGAGAGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAAAAGTATCATG
                ************************************************************

TaSUT2A         ACGGTGTGAGAATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCTCGGGATCACAT
TaSUT2B         ACGGTGTGAGAATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCTCGGGATCACAT
TaSUT2D         ACGGTGTGAGAATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCTCGGGATCACAT
                ************************************************************

TaSUT2A         CTATTGGAATGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTATGGGGTGTCTCCA
TaSUT2B         CTATTGGAATGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTATGGGGTGTCTCCA
TaSUT2D         CTATTGGAATGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTATGGGGTGTCTCCA
                ************************************************************

TaSUT2A         ATATCATCATGGCACTGTGCTTCGTGGCGATGCTTATTATAACATACGTTGCGCAGAACC
TaSUT2B         ATATCATCATGGCACTGTGCTTCGTGGCGATGCTTATTATAACATACGTTGCGCAGAACC
TaSUT2D         ATATCATCATGGCACTGTGCTTCGTGGCGATGCTTATTATAACATACGTTGCGCAGAACC
                ************************************************************

TaSUT2A         TGGATTATGGACCTAGTGGAGCACCTCCAACCGGCATTGTCGCCGCTTCCCTCATTGTGT
TaSUT2B         TGGATTATGGACCTAGTGGAGCACCTCCAACCGACATTGTCGCCGCTTCCCTCATTGTGT
TaSUT2D         TGGATTATGGACCTAGTGGAGCACCTCCAACCGGCATTGTCGCCGCTTCCCTCATTGTGT
                ********************************* ************************

TaSUT2A         TCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTATACCATATGCGATGGCTGCAA
TaSUT2B         TCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTATACCATATGCGATGGCTGCAA
TaSUT2D         TCACAATCCTAGGAGCACCTCTCTCGGTCACGTACAGTATACCATATGCGATGGCTGCAA
                ********************** ***** ********************************

TaSUT2A         GTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGCATTCTTAATTTATCTA
TaSUT2B         GTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGCATTCTTAATTTATCTA
TaSUT2D         GTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGCATTCTTAATTTATCTA
                ************************************************************

TaSUT2A         TTGTCATACCACAGATCATCGTGTCGCTGGGCAGCGGGCCATGGGACCAGCTCTTCGGCG
TaSUT2B         TTGTCATACCACAGATCATCGTGTCGCTGGGCAGCGGGCCATGGGACCAGCTCTTCGGCG
TaSUT2D         TTGTCATACCACAGATCATCGTGTCGCTGGGCAGCGGGCCATGGGACCAGCTCTTCGGCG
                ************************************************************

TaSUT2A         GAGGGAACGCGCCATCCTTCTGGGTGGCTGCCGCAGCCTCCTTTGTGGGCGGGCTGGTAG
TaSUT2B         GAGGGAACGCGCCATCCTTCTGGGTGGCTGCCGCAGCCTCCTTTGTGGGCGGGCTGGTAG
TaSUT2D         GAGGGAACGCGCCATCCTTCTGGGTGGCTGCCGCAGCCTCCTTTGTGGGCGGGCTGGTAG
                ************************************************************

TaSUT2A         CCATCCTGGGGCTCCCGCGGGCCCGGCTCGGGCCGAAGAAGAAAACCACCCAACGATGAT
TaSUT2B         CCATCCTGGGGCTCCCGCGGGCCCGGCTCGGGCCGAAGAAGAAAACCACCCAACGATGAT
TaSUT2D         CCATCCTGGGGCTCCCGCGGGCCCGGCTCGGGCCGAAGAAGAAAACCACCCAACGATGAT
                ************************************************************

TaSUT2A         GATTCGAGCGAGTATACAGAGACATTGTTATGTTGCAGACAAGGAAGGAAGGCAGGTTGT
TaSUT2B         GATTCGAGCGAGTATACAGAGACATTGTTATGTTGCAGACAAGGAAGGAAGGCAGGTTGT
TaSUT2D         GATTCGAGCGAGTATACAGAGACATTGTTATGTTGCAGACAAGGAAGGAAGGCAGGTTGT
                ************************************************************
```

Continued

```
TaSUT2A         ACACTGGTTCATTCTTTGCACACTGGTTCCAAGACACCGGGAGCGCTGCCTACACTACAC
TaSUT2B         ACACTGGTTCATTCTTTGCACACTGGTTCCAAGACACCGGGAGCGCTGCCTACACTACAC
TaSUT2D         ACACTGGTTCATTCTTTGCACACTGGTTCCAAGACACCGGGAACGCTGCCTACACTACAC
                **************************************** ****************

TaSUT2A         TAGGCATGGGGGGTCTGTATATTGTCATCATCAAGTTGGATTGGGTTCAAATTCAAGGCC
TaSUT2B         TAGGCATGGGGGGTCTGTATATTGTCATCATCAAGTTGGATTGGGTTCAAATTCAAGGCC
TaSUT2D         TAGGCATGGGGGGTCTGTATATTGTCATCATCAAGTTGGATTGGGTTCAAATTCAAGGCC
                ************************************************************

TaSUT2A         AAAAGTAGGAGGGCTCTGGAATTTTTCTTCCTCTTGTTACTTTCCTCCATGGGCCCTATC
TaSUT2B         AAAAGTAGGAGGGCTCTGGAAT-----------------TTTCCTCCATGGGCCCTAGC
TaSUT2D         AAAAGTAGGAGGGCTCTGGAATTTTTCTTCCTCTTGTTACTTTCCTCCATGGGCCCTATC
                ********************                  ****************** *

TaSUT2A         ATTGCCTTTTTTTCGCTTGTCGGTGTGAAACATGGTCAGTCATATGTTGTTGTATATTCG
TaSUT2B         GTCGCCCTTTTTCCGCTTGTCGGTGTGAAAAATGGTCAGTCATATGTTGTTGTATATT-G
TaSUT2D         ATCGCCTTTTTTTCGCTTGTCGGTGTGAAACATGGTCAGTCATATGTTGTTGTATATTCG
                 * *** ***** ***************** *************************** *

TaSUT2A         GATATTTTGTCATTTTGTCGGAGATTATCAATCAATCAATAAAAGTTGCAGGTTTTTGTT
TaSUT2B         GATATTTTGCCATTTTGTTGGAGATTATCAACCAATCAATAAAAGTTGCAGGTTTTTGTT
TaSUT2D         GATATTTTGTCATTTTGTCGGAGATTATCAATCAATCAATAAAAGTTGCAGGTTTTTGTT
                ********* ******** ************ ****************************

TaSUT2A         TGAAA
TaSUT2B         TGAAA
TaSUT2D         TGAAA
                *****
```

Figure 3.19 Alignment of the nucleotide sequences of *SUT2* identified from the three diploid progenitors of hexaploid wheat; *Triticum urartu* (A genome donor), *Aegilops speltoides* (B genome donor), *Aegilops tauschii* (D genome donor). The coding sequence of each gene lies between the bolded and underlined start (ATG) and stop (TGA) codons.

```
TaSUT2A         MPPRRPNTGAGGGGTSSAA---PPRKVPLRSLLRAASVACGVQFGWALQLSLLTPYVQEL
TaSUT2B         MPPRRPNTGAGGGGTSSAA---PPRKVPLRSLLRAASVACGVQFGWALQLSLLTPYVQEL
TaSUT2D         MPPRRPNTG-GGGGTSSAAPPPPPRKVPLRSLLRAASVACGVQFGWALQLSLLTPYVQEL
                ********* *********   ************************************

TaSUT2A         GIPHAFASLVWLCGPLSGLLVQPLVGHLSDRIAPANSPLGRRRPFIAAGAASIAFSVLTV
TaSUT2B         GIPHAFASLVWLCGPLSGLLVQPLVGHLSDRIAPANSPLGRRRPFIAAGAASIAFSVLTV
TaSUT2D         GIPHAFASLVWLCGPLSGLLVQPLVGHLSDRIAPANSPLGRRRPFIAAGAASIAFSVLTV
                ************************************************************

TaSUT2A         GFSADLGRLFGDNIVPGSTRFGAIIVYLIGFWLLDVGNNATQGPCRAFLADLTENDPRRT
TaSUT2B         GFSADLGRLFGDNIVPGSTRFGAIIVYLIGFWLLDVGNNATQGPCRAFLADLTENDPRRT
TaSUT2D         GFSADLGRLFGDNIVPGSTRFGAIIVYLIGFWLLDVGNNATQGPCRAFLADLTENDPRRT
                ************************************************************

TaSUT2A         RIANAYFSLFMALGNILGYATGAYNGWYKIFPFTITESCSVSCANLKSAFLLDIIILAIT
TaSUT2B         RIANAYFSLFMALGNILGYATGAYNGWYKIFPFTITESCSVSCANLKSAFLLDIIILAIT
TaSUT2D         RIANAYFSLFMALGNILGYATGAYNGWYKIFPFTITESCSVSCANLKSAFLLDIIILAIT
                ************************************************************

TaSUT2A         TYVSVVTVQDNPTFGSDEAAPPSSHEEEAFLFELFGSFKYFTMPVWMVLIVTSLTWIGWF
TaSUT2B         TYVSVVTVQDNPTFGSDEAAPPSSHEEEAFLFELFGSFKYFTMPVWMVLIVTSLTWIGWF
TaSUT2D         TYVSVVTVQDNPTFGSDEAAPPSSHEEEAFLFELFGSFKYFTMPVWMVLIVTSLTWIGWF
                ************************************************************

TaSUT2A         PFILFDTDWMGREIYRGSPEIVADTQKYHDGVRMGSFGLMLNSVVLGITSIGMEKLCRKW
TaSUT2B         PFILFDTDWMGREIYRGSPEIVADTQKYHDGVRMGSFGLMLNSVVLGITSIGMEKLCRKW
TaSUT2D         PFILFDTDWMGREIYRGSPEIVADTQKYHDGVRMGSFGLMLNSVVLGITSIGMEKLCRKW
                ************************************************************

TaSUT2A         GAGLVWGVSNIIMALCFVAMLIITYVAQNLDYGPSGAPPTGIVAASLIVFTILGAPLSVT
TaSUT2B         GAGLVWGVSNIIMALCFVAMLIITYVAQNLDYGPSGAPPTDIVAASLIVFTILGAPLSVT
TaSUT2D         GAGLVWGVSNIIMALCFVAMLIITYVAQNLDYGPSGAPPTGIVAASLIVFTILGAPLSVT
                ************************************ ******************

TaSUT2A         YSIPYAMAASRVENLGLGQGLAMGILNLSIVIPQIIVSLGSGPWDQLFGGGNAPSFWVAA
TaSUT2B         YSIPYAMAASRVENLGLGQGLAMGILNLSIVIPQIIVSLGSGPWDQLFGGGNAPSFWVAA
TaSUT2D         YSIPYAMAASRVENLGLGQGLAMGILNLSIVIPQIIVSLGSGPWDQLFGGGNAPSFWVAA
                ************************************************************

TaSUT2A         AASFVGGLVAILGLPRARLGPKKKTTQRX
TaSUT2B         AASFVGGLVAILGLPRARLGPKKKTTQRX
TaSUT2D         AASFVGGLVAILGLPRARLGPKKKTTQRX
                ****************************
```

Figure 3.20 Alignment of the amino acid sequences of SUT2 from the three diploid progenitors of hexaploid wheat; *Triticum urartu* (A genome donor), *Aegilops speltoides* (B genome donor), *Aegilops tauschii* (D genome donor).

3.3.3 Identification of *TaSUT2* from the A, B and D genomes of hexaploid wheat cv. AC Andrew

Sequencing of multiple colonies obtained by transformation of *E.coli* (DH5α) cells with *TaSUT2* derived from cv. AC Andrew by RACE-PCR and subsequent comparison of the resulting cDNA sequences with those obtained from the three diploid progenitors enabled us to identify the three homeologues of *SUT2* derived from the A, B and D genome of hexaploid wheat. Two different cDNA sequences derived from AC Andrew showed 100% similarity, one with the SUT2 cDNA sequence derived from *Triticum urartu* (Figure 3.21) and the other one with that obtained from *Aegilops tauschii* (Figure 3.22). The third cDNA sequence from AC Andrew showed 99.5% similarity with that derived from *Aegilops speltoides* with differences in one base in the coding region and few bases in the 3' untranslated region (Fig 3.23).

```
TuSUT2    ACGCGGGGAGAGCGATAGATACTTTGCTTTCCCCTCCCTCCATCCCTTCCCTTCCCTCTC
TaSUT2A   ACGCGGGGAGAGCGATAGATACTTTGCTTTCCCCTCCCTCCATCCCTTCCCTTCCCTCTC
          ************************************************************

TuSUT2    TCCCAGGTCAGGCCAGGTCTCCCACAGAATTCAGCCAGGAGGCGTGGGGGGAGCGGGCAG
TaSUT2A   TCCCAGGTCAGGCCAGGTCTCCCACAGAATTCAGCCAGGAGGCGTGGGGGGAGCGGGCAG
          ************************************************************

TuSUT2    CCAAGAAAGTCCAACACCACTCATCAAGGGCGATGCCGCCGCGGCGGCCCAACACCGGCG
TaSUT2A   CCAAGAAAGTCCAACACCACTCATCAAGGGCGATGCCGCCGCGGCGGCCCAACACCGGCG
          ************************************************************

TuSUT2    CCGGGGGCGGGGGCACCTCCTCGGCCGCGCCGCCGCGCAAGGTCCCGCTCCGCTCGCTGC
TaSUT2A   CCGGGGGCGGGGGCACCTCCTCGGCCGCGCCGCCGCGCAAGGTCCCGCTCCGCTCGCTGC
          ************************************************************

TuSUT2    TCCGGGCGGCGTCCGTCGCCTGCGGGGTCCAATTCGGGTGGGCGCTGCAGCTCTCCCTGC
TaSUT2A   TCCGGGCGGCGTCCGTCGCCTGCGGGGTCCAATTCGGGTGGGCGCTGCAGCTCTCCCTGC
          ************************************************************

TuSUT2    TCACCCCGTACGTCCAAGAACTCGGCATCCCGCACGCCTTCGCCAGCCTGGTCTGGCTCT
TaSUT2A   TCACCCCGTACGTCCAAGAACTCGGCATCCCGCACGCCTTCGCCAGCCTGGTCTGGCTCT
          ************************************************************
```

Continued

```
TuSUT2   GCGGCCCGCTGTCGGGCCTCCTCGTGCAGCCCCTGGTGGGCCACCTCTCGGACCGCATCG
TaSUT2A  GCGGCCCGCTGTCGGGCCTCCTCGTGCAGCCCCTGGTGGGCCACCTCTCGGACCGCATCG
         ************************************************************

TuSUT2   CGCCGGCCAACTCCCCGCTGGGCCGCCGCCGGCCCTTCATCGCCGCGGGGGCCGCCTCCA
TaSUT2A  CGCCGGCCAACTCCCCGCTGGGCCGCCGCCGGCCCTTCATCGCCGCGGGGGCCGCCTCCA
         ************************************************************

TuSUT2   TCGCCTTCTCCGTGCTCACCGTCGGCTTCTCCGCCGACCTCGGCCGCCTCTTCGGGGACA
TaSUT2A  TCGCCTTCTCCGTGCTCACCGTCGGCTTCTCCGCCGACCTCGGCCGCCTCTTCGGGGACA
         ************************************************************

TuSUT2   ACATCGTGCCCGGGTCCACCCGGTTCGGCGCCATCATCGTCTACCTCATCGGCTTCTGGC
TaSUT2A  ACATCGTGCCCGGGTCCACCCGGTTCGGCGCCATCATCGTCTACCTCATCGGCTTCTGGC
         ************************************************************

TuSUT2   TGCTGGATGTCGGCAACAACGCCACCCAGGGGCCATGCCGCGCCTTCCTCGCAGACCTCA
TaSUT2A  TGCTGGATGTCGGCAACAACGCCACCCAGGGGCCATGCCGCGCCTTCCTCGCAGACCTCA
         ************************************************************

TuSUT2   CAGAGAATGACCCGAGGAGGACCCGGATTGCCAACGCCTACTTCTCGCTCTTCATGGCCC
TaSUT2A  CAGAGAATGACCCGAGGAGGACCCGGATTGCCAACGCCTACTTCTCGCTCTTCATGGCCC
         ************************************************************

TuSUT2   TGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGATATTCCCCT
TaSUT2A  TGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGATATTCCCCT
         ************************************************************

TuSUT2   TCACTATTACGGAGTCCTGCTCTGTCAGCTGCGCCAACCTCAAGTCCGCGTTCCTGCTCG
TaSUT2A  TCACTATTACGGAGTCCTGCTCTGTCAGCTGCGCCAACCTCAAGTCCGCGTTCCTGCTCG
         ************************************************************

TuSUT2   ACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGTGACGGTGCAGGACAATCCTA
TaSUT2A  ACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGTGACGGTGCAGGACAATCCTA
         ************************************************************

TuSUT2   CTTTCGGAAGCGACGAGGCGGCACCTCCGAGCAGCCACGAGGAGGAGGCTTTCCTCTTTG
TaSUT2A  CTTTCGGAAGCGACGAGGCGGCACCTCCGAGCAGCCACGAGGAGGAGGCTTTCCTCTTTG
         ************************************************************

TuSUT2   AGCTTTTCGGGTCGTTCAAGTACTTCACGATGCCTGTTTGGATGGTCTTGATCGTCACCT
TaSUT2A  AGCTTTTCGGGTCGTTCAAGTACTTCACGATGCCTGTTTGGATGGTCTTGATCGTCACCT
         ************************************************************

TuSUT2   CGCTTACCTGGATCGGCTGGTTCCCTTTCATCCTCTTTGACACCGACTGGATGGGCCGAG
TaSUT2A  CGCTTACCTGGATCGGCTGGTTCCCTTTCATCCTCTTTGACACCGACTGGATGGGCCGAG
         ************************************************************

TuSUT2   AGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAAAAGTATCATGACGGTGTGA
TaSUT2A  AGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAAAAGTATCATGACGGTGTGA
         ************************************************************

TuSUT2   GAATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCTCGGGATCACATCTATTGGAA
TaSUT2A  GAATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCTCGGGATCACATCTATTGGAA
         ************************************************************

TuSUT2   TGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTATGGGGTGTCTCCAATATCATCA
TaSUT2A  TGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTATGGGGTGTCTCCAATATCATCA
         ************************************************************
```

Continued

```
TuSUT2    TGGCACTGTGCTTCGTGGCGATGCTTATTATAACATACGTTGCGCAGAACCTGGATTATG
TaSUT2A   TGGCACTGTGCTTCGTGGCGATGCTTATTATAACATACGTTGCGCAGAACCTGGATTATG
          ************************************************************

TuSUT2    GACCTAGTGGAGCACCTCCAACCGGCATTGTCGCCGCTTCCCTCATTGTGTTCACAATCC
TaSUT2A   GACCTAGTGGAGCACCTCCAACCGGCATTGTCGCCGCTTCCCTCATTGTGTTCACAATCC
          ************************************************************

TuSUT2    TAGGAGCACCTCTGTCGGTTACGTACAGTATACCATATGCGATGGCTGCAAGTCGTGTTG
TaSUT2A   TAGGAGCACCTCTGTCGGTTACGTACAGTATACCATATGCGATGGCTGCAAGTCGTGTTG
          ************************************************************

TuSUT2    AAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGCATTCTTAATTTATCTATTGTCATAC
TaSUT2A   AAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGCATTCTTAATTTATCTATTGTCATAC
          ************************************************************

TuSUT2    CACAGATCATCGTGTCGCTGGGCAGCGGGCCATGGGACCAGCTCTTCGGCGGAGGGAACG
TaSUT2A   CACAGATCATCGTGTCGCTGGGCAGCGGGCCATGGGACCAGCTCTTCGGCGGAGGGAACG
          ************************************************************

TuSUT2    CGCCATCCTTCTGGGTGGCTGCCGCAGCCTCCTTTGTGGGCGGGCTGGTAGCCATCCTGG
TaSUT2A   CGCCATCCTTCTGGGTGGCTGCCGCAGCCTCCTTTGTGGGCGGGCTGGTAGCCATCCTGG
          ************************************************************

TuSUT2    GGCTCCCGCGGGCCCGGCTCGGGCCGAAGAAGAAAACCACCCAACGATGATGATTCGAGC
TaSUT2A   GGCTCCCGCGGGCCCGGCTCGGGCCGAAGAAGAAAACCACCCAACGATGATGATTCGAGC
          ************************************************************

TuSUT2    GAGTATACAGAGACATTGTTATGTTGCAGACAAGGAAGGAAGGCAGGTTGTACACTGGTT
TaSUT2A   GAGTATACAGAGACATTGTTATGTTGCAGACAAGGAAGGAAGGCAGGTTGTACACTGGTT
          ************************************************************

TuSUT2    CATTCTTTGCACACTGGTTCCAAGACACCGGGAGCGCTGCCTACACTACACTAGGCATGG
TaSUT2A   CATTCTTTGCACACTGGTTCCAAGACACCGGGAGCGCTGCCTACACTACACTAGGCATGG
          ************************************************************

TuSUT2    GGGGTCTGTATATTGTCATCATCAAGTTGGATTGGGTTCAAATTCAAGGCCAAAAGTAGG
TaSUT2A   GGGGTCTGTATATTGTCATCATCAAGTTGGATTGGGTTCAAATTCAAGGCCAAAAGTAGG
          ************************************************************

TuSUT2    AGGGCTCTGGAATTTTTCTTCCTCTTGTTACTTTCCTCCATGGGCCCTATCATTGCCTTT
TaSUT2A   AGGGCTCTGGAATTTTTCTTCCTCTTGTTACTTTCCTCCATGGGCCCTATCATTGCCTTT
          ************************************************************

TuSUT2    TTTTCGCTTGTCGGTGTGAAACATGGTCAGTCATATGTTGTTGTATATTCGGATATTTTG
TaSUT2A   TTTTCGCTTGTCGGTGTGAAACATGGTCAGTCATATGTTGTTGTATATTCGGATATTTTG
          ************************************************************

TuSUT2    TCATTTTGTCGGAGATTATCAATCAATCAATAAAAGTTGCAGGTTTTTGTTTGAAA
TaSUT2A   TCATTTTGTCGGAGATTATCAATCAATCAATAAAAGTTGCAGGTTTTTGTTTGAAA
          ***************************************************
```

Figure 3.21 Alignment of the nucleotide sequences of *TaSUT2* identified from *Triticum urartu* (A genome donor; *TuSUT2*) and *Triticum aestivum* cv. AC Andrew *(TaSUT2A).*

```
AetSUT2   ACGCGGGGAGAGCGATAGATACTTTGCTTTCCCCTCCCTCCATCCCTTCCCTTCCCTCTC
TaSUT2D   ACGCGGGGAGAGCGATAGATACTTTGCTTTCCCCTCCCTCCATCCCTTCCCTTCCCTCTC
          ************************************************************

AetSUT2   TCCCAGGTCAGGCCAGGTCTCCCGCAGAATTCAGCCAGGAGGCGTGGGGGGAGCGGGCAG
TaSUT2D   TCCCAGGTCAGGCCAGGTCTCCCGCAGAATTCAGCCAGGAGGCGTGGGGGGAGCGGGCAG
          ************************************************************

AetSUT2   CCAAGAAAGTCCAACACCGCTCATCAAGGGCGATGCCGCCGCGCCGGCCCAACACCGGCG
TaSUT2D   CCAAGAAAGTCCAACACCGCTCATCAAGGGCGATGCCGCCGCGCCGGCCCAACACCGGCG
          ************************************************************

AetSUT2   GGGGCGGGGGCACCTCCTCGGCCGCGCCGCCGCCTCCGCCGCGCAAGGTCCCGCTCCGCT
TaSUT2D   GGGGCGGGGGCACCTCCTCGGCCGCGCCGCCGCCTCCGCCGCGCAAGGTCCCGCTCCGCT
          ************************************************************

AetSUT2   CGCTGCTCCGGGCGGCCTCCGTCGCCTGCGGGGTCCAATTCGGGTGGGCGCTGCAGCTCT
TaSUT2D   CGCTGCTCCGGGCGGCCTCCGTCGCCTGCGGGGTCCAATTCGGGTGGGCGCTGCAGCTCT
          ************************************************************

AetSUT2   CCCTGCTCACCCCCTACGTCCAAGAGCTCGGCATCCCGCACGCCTTCGCCAGCCTGGTCT
TaSUT2D   CCCTGCTCACCCCCTACGTCCAAGAGCTCGGCATCCCGCACGCCTTCGCCAGCCTGGTCT
          ************************************************************

AetSUT2   GGCTCTGCGGCCCGCTGTCGGGCCTCCTCGTGCAGCCCCTGGTGGGCCACCTCTCGGACC
TaSUT2D   GGCTCTGCGGCCCGCTGTCGGGCCTCCTCGTGCAGCCCCTGGTGGGCCACCTCTCGGACC
          ************************************************************

AetSUT2   GCATCGCGCCGGCCAACTCGCCGCTCGGCCGCCGCCGGCCCTTCATCGCCGCGGGGGCCG
TaSUT2D   GCATCGCGCCGGCCAACTCGCCGCTCGGCCGCCGCCGGCCCTTCATCGCCGCGGGGGCCG
          ************************************************************

AetSUT2   CCTCCATCGCCTTCTCCGTGCTCACCGTCGGCTTCTCCGCCGACCTGGGCCGCCTCTTCG
TaSUT2D   CCTCCATCGCCTTCTCCGTGCTCACCGTCGGCTTCTCCGCCGACCTGGGCCGCCTCTTCG
          ************************************************************

AetSUT2   GGGACAACATCGTGCCCGGGTCCACCCGGTTCGGCGCCATCATCGTCTACCTCATCGGCT
TaSUT2D   GGGACAACATCGTGCCCGGGTCCACCCGGTTCGGCGCCATCATCGTCTACCTCATCGGCT
          ************************************************************

AetSUT2   TCTGGCTGCTGGATGTGGGCAACAACGCCACCCAGGGGCCATGCCGCGCCTTCCTCGCAG
TaSUT2D   TCTGGCTGCTGGATGTGGGCAACAACGCCACCCAGGGGCCATGCCGCGCCTTCCTCGCAG
          ************************************************************

AetSUT2   ACCTCACAGAGAATGACCCGAGGAGGACCCGGATTGCCAACGCCTACTTCTCGCTCTTCA
TaSUT2D   ACCTCACAGAGAATGACCCGAGGAGGACCCGGATTGCCAACGCCTACTTCTCGCTCTTCA
          ************************************************************

AetSUT2   TGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGATAT
TaSUT2D   TGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGATAT
          ************************************************************

AetSUT2   TCCCCTTCACTATTACCGAGTCCTGCTCTGTCAGCTGCGCCAACCTCAAGTCCGCGTTCC
TaSUT2D   TCCCCTTCACTATTACCGAGTCCTGCTCTGTCAGCTGCGCCAACCTCAAGTCCGCGTTCC
          ************************************************************

AetSUT2   TGCTCGACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGTGACGGTGCAGGACA
TaSUT2D   TGCTCGACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGTGACGGTGCAGGACA
          ************************************************************
```

Continued

```
AetSUT2   ATCCTACTTTCGGAAGCGACGAGGCGGCGCCTCCGAGCAGCCACGAGGAGGAGGCTTTCC
TaSUT2D   ATCCTACTTTCGGAAGCGACGAGGCGGCGCCTCCGAGCAGCCACGAGGAGGAGGCTTTCC
          ************************************************************

AetSUT2   TCTTTGAGCTTTTCGGGTCGTTCAAGTACTTCACGATGCCTGTTTGGATGGTCTTGATCG
TaSUT2D   TCTTTGAGCTTTTCGGGTCGTTCAAGTACTTCACGATGCCTGTTTGGATGGTCTTGATCG
          ************************************************************

AetSUT2   TCACCTCGCTTACCTGGATCGGCTGGTTCCCTTTCATCCTCTTTGACACCGACTGGATGG
TaSUT2D   TCACCTCGCTTACCTGGATCGGCTGGTTCCCTTTCATCCTCTTTGACACCGACTGGATGG
          ************************************************************
AetSUT2   GCCGAGAGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAAAAGTATCATGACG
TaSUT2D   GCCGAGAGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAAAAGTATCATGACG
          ************************************************************

AetSUT2   GTGTGAGAATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCTCGGGATCACATCTA
TaSUT2D   GTGTGAGAATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCTCGGGATCACATCTA
          ************************************************************

AetSUT2   TTGGAATGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTATGGGGTGTCTCCAATA
TaSUT2D   TTGGAATGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTATGGGGTGTCTCCAATA
          ************************************************************

AetSUT2   TCATCATGGCACTGTGCTTCGTGGCGATGCTTATTATAACATACGTTGCGCAGAACCTGG
TaSUT2D   TCATCATGGCACTGTGCTTCGTGGCGATGCTTATTATAACATACGTTGCGCAGAACCTGG
          ************************************************************

AetSUT2   ATTATGGACCTAGTGGAGCACCTCCAACCGGCATTGTCGCCGCTTCCCTCATTGTGTTCA
TaSUT2D   ATTATGGACCTAGTGGAGCACCTCCAACCGGCATTGTCGCCGCTTCCCTCATTGTGTTCA
          ************************************************************

AetSUT2   CAATCCTAGGAGCACCTCTCTCGGTCACGTACAGTATACCATATGCGATGGCTGCAAGTC
TaSUT2D   CAATCCTAGGAGCACCTCTCTCGGTCACGTACAGTATACCATATGCGATGGCTGCAAGTC
          ************************************************************

AetSUT2   GTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGCATTCTTAATTTATCTATTG
TaSUT2D   GTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGCATTCTTAATTTATCTATTG
          ************************************************************

AetSUT2   TCATACCACAGATCATCGTGTCGCTGGGCAGCGGGCCATGGGACCAGCTCTTCGGCGGAG
TaSUT2D   TCATACCACAGATCATCGTGTCGCTGGGCAGCGGGCCATGGGACCAGCTCTTCGGCGGAG
          ************************************************************

AetSUT2   GGAACGCGCCATCCTTCTGGGTGGCTGCCGCAGCCTCCTTTGTGGGCGGGCTGGTAGCCA
TaSUT2D   GGAACGCGCCATCCTTCTGGGTGGCTGCCGCAGCCTCCTTTGTGGGCGGGCTGGTAGCCA
          ************************************************************

AetSUT2   TCCTGGGGCTCCCGCGGGCCCGGCTCGGGCCGAAGAAGAAAACCACCCAACGATGATGAT
TaSUT2D   TCCTGGGGCTCCCGCGGGCCCGGCTCGGGCCGAAGAAGAAAACCACCCAACGATGATGAT
          ************************************************************

AetSUT2   TCGAGCGAGTATACAGAGACATTGTTATGTTGCAGACAAGGAAGGAAGGCAGGTTGTACA
TaSUT2D   TCGAGCGAGTATACAGAGACATTGTTATGTTGCAGACAAGGAAGGAAGGCAGGTTGTACA
          ************************************************************

AetSUT2   CTGGTTCATTCTTTGCACACTGGTTCCAAGACACCGGGAACGCTGCCTACACTACACTAG
TaSUT2D   CTGGTTCATTCTTTGCACACTGGTTCCAAGACACCGGGAACGCTGCCTACACTACACTAG
          ************************************************************
```

Continued

```
AetSUT2   GCATGGGGGGTCTGTATATTGTCATCATCAAGTTGGATTGGGTTCAAATTCAAGGCCAAA
TaSUT2D   GCATGGGGGGTCTGTATATTGTCATCATCAAGTTGGATTGGGTTCAAATTCAAGGCCAAA
          ************************************************************

AetSUT2   AGTAGGAGGGCTCTGGAATTTTTCTTCCTCTTGTTACTTTCCTCCATGGGCCCTATCATC
TaSUT2D   AGTAGGAGGGCTCTGGAATTTTTCTTCCTCTTGTTACTTTCCTCCATGGGCCCTATCATC
          ************************************************************

AetSUT2   GCCTTTTTTTCGCTTGTCGGTGTGAAACATGGTCAGTCATATGTTGTTGTATATTCGGAT
TaSUT2D   GCCTTTTTTTCGCTTGTCGGTGTGAAACATGGTCAGTCATATGTTGTTGTATATTCGGAT
          ************************************************************

AetSUT2   ATTTTGTCATTTTGTCGGAGATTATCAATCAATCAATAAAAGTTGCAGGTTTTTGTTTGA
TaSUT2D   ATTTTGTCATTTTGTCGGAGATTATCAATCAATCAATAAAAGTTGCAGGTTTTTGTTTGA
          ************************************************************

AetSUT2   AA
TaSUT2D   AA
          **
```

Figure 3.22 Alignment of the nucleotide sequences of *TaSUT2* identified from *Aegilops tauschii* (D genome donor; *AetSUT2*) and *Triticum aestivum* cv. AC Andrew (*TaSUT2D*).

```
AesSUT2   ACGCGGGGAGAGCGATAGATACTTTGCTTTCCCCTCCCTCCATCCCTTCCCTTCCCTCTC
TaSUT2B   ACGCGGGGAGAGCGATAGATACTTTGCTTTCCCCTCCCTCCATCCCTTCCCTTCCCTCTC
          ************************************************************

AesSUT2   TCCCAGGTCAGGCCAGGTCTCCCACAGAATTCAGCCAGGAGGCGTGGGGGGAGCGGGCAG
TaSUT2B   TCCCAGGTCAGGCCAGGTCTCCCACAGAATTCAGCCAGGAGGCGTGGGGGGAGCGGGCAG
          ************************************************************

AesSUT2   CCAAGAAAGTCCAACACCACTCATCAAGGGCGATGCCGCCGCGGCGGCCCAACACCGGCG
TaSUT2B   CCAAGAAAGTCCAACACCACTCATCAAGGGCGATGCCGCCGCGGCGGCCCAACACCGGCG
          ************************************************************

AesSUT2   CCGGGGGCGGGGGCACCTCCTCGGCCGCGCCGCCGCGCAAGGTCCCGCTCCGCTCGCTGC
TaSUT2B   CCGGGGGCGGGGGCACCTCCTCGGCCGCGCCGCCGCGCAAGGTCCCGCTCCGCTCGCTGC
          ************************************************************

AesSUT2   TCCGGGCGGCGTCCGTCGCCTGCGGGGTCCAATTCGGGTGGGCGCTGCAGCTCTCCCTGC
TaSUT2B   TCCGGGCGGCGTCCGTCGCCTGCGGGGTCCAATTCGGGTGGGCGCTGCAGCTCTCCCTGC
          ************************************************************

AesSUT2   TCACCCCGTACGTCCAAGAACTCGGCATCCCGCACGCCTTCGCCAGCCTGGTCTGGCTCT
TaSUT2B   TCACCCCGTACGTCCAAGAACTCGGCATCCCGCACGCCTTCGCCAGCCTGGTCTGGCTCT
          ************************************************************

AesSUT2   GCGGCCCGCTGTCGGGCCTCCTCGTGCAGCCCCTGGTGGGCCACCTCTCGGACCGCATCG
TaSUT2B   GCGGCCCGCTGTCGGGCCTCCTCGTGCAGCCCCTGGTGGGCCACCTCTCGGACCGCATCG
          ************************************************************

AesSUT2   CGCCGGCCAACTCCCCGCTGGGCCGCCGCCGGCCCTTCATCGCCGCGGGGGCCGCCTCCA
TaSUT2B   CGCCGGCCAACTCCCCGCTGGGCCGCCGCCGGCCCTTCATCGCCGCGGGGGCCGCCTCCA
          ************************************************************
```

Continued

```
AesSUT2   TCGCCTTCTCCGTGCTCACCGTCGGCTTCTCCGCCGACCTCGGCCGCCTCTTCGGGGACA
TaSUT2B   TCGCCTTCTCCGTGCTCACCGTCGGCTTCTCCGCCGACCTCGGCCGCCTCTTCGGGGACA
          ************************************************************

AesSUT2   ACATCGTGCCCGGGTCCACCCGGTTCGGCGCCATCATCGTCTACCTCATCGGCTTCTGGC
TaSUT2B   ACATCGTGCCCGGGTCCACCCGGTTCGGCGCCATCATCGTCTACCTCATCGGCTTCTGGC
          ************************************************************

AesSUT2   TGCTGGATGTCGGCAACAACGCCACCCAGGGGCCATGCCGCGCCTTCCTCGCAGACCTCA
TaSUT2B   TGCTGGATGTCGGCAACAACGCCACCCAGGGGCCATGCCGCGCCTTCCTCGCAGACCTCA
          ************************************************************

AesSUT2   CAGAGAATGACCCGAGGAGGACCCGGATTGCCAACGCCTACTTCTCGCTCTTCATGGCCC
TaSUT2B   CAGAGAATGACCCGAGGAGGACCCGGATTGCCAACGCCTACTTCTCGCTCTTCATGGCCC
          ************************************************************

AesSUT2   TGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGATATTCCCCT
TaSUT2B   TGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGATATTCCCCT
          ************************************************************

AesSUT2   TCACTATTACGGAGTCCTGCTCTGTCAGCTGCGCCAACCTCAAGTCCGCGTTCCTGCTCG
TaSUT2B   TCACTATTACGGAGTCCTGCTCTGTCAGCTGCGCCAACCTCAAGTCCGCGTTCCTGCTCG
          ************************************************************

AesSUT2   ACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGTGACGGTGCAGGACAATCCTA
TaSUT2B   ACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGTGACGGTGCAGGACAATCCTA
          ************************************************************

AesSUT2   CTTTCGGAAGCGACGAGGCGGCACCTCCGAGCAGCCACGAGGAGGAGGCTTTCCTCTTTG
TaSUT2B   CTTTCGGAAGCGACGAGGCGGCACCTCCGAGCAGCCACGAGGAGGAGGCTTTCCTCTTTG
          ************************************************************
AesSUT2   AGCTTTTCGGGTCGTTCAAGTACTTCACGATGCCTGTTTGGATGGTCTTGATCGTCACCT
TaSUT2B   AGCTTTTCGGGTCGTTCAAGTACTTCACGATGCCTGTTTGGATGGTCTTGATCGTCACCT
          ************************************************************

AesSUT2   CGCTTACCTGGATCGGCTGGTTCCCTTTCATCCTCTTTGACACCGACTGGATGGGCCGAG
TaSUT2B   CGCTTACCTGGATCGGCTGGTTCCCTTTCATCCTCTTTGACACCGACTGGATGGGCCGAG
          ************************************************************

AesSUT2   AGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAAAAGTATCATGACGGTGTGA
TaSUT2B   AGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAAAAGTATCATGACGGTGTGA
          ************************************************************

AesSUT2   GAATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCTCGGGATCACATCTATTGGAA
TaSUT2B   GAATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCTCGGGATCACATCTATTGGAA
          ************************************************************

AesSUT2   TGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTATGGGGTGTCTCCAATATCATCA
TaSUT2B   TGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTATGGGGTGTCTCCAATATCATCA
          ************************************************************

AesSUT2   TGGCACTGTGCTTCGTGGCGATGCTTATTATAACATACGTTGCGCAGAACCTGGATTATG
TaSUT2B   TGGCACTGTGCTTCGTGGCGATGCTTATTATAACATACGTTGCGCAGAACCTGGATTATG
          ************************************************************

AesSUT2   GACCTAGTGGAGCACCTCCAACCGACATTGTCGCCGCTTCCCTCATTGTGTTCACAATCC
TaSUT2B   GACCTAGTGGAGCACCTCCAACCGGCATTGTCGCCGCTTCCCTCATTGTGTTCACAATCC
          ************************ ************************************
```

Continued

```
AesSUT2   TAGGAGCACCTCTGTCGGTTACGTACAGTATACCATATGCGATGGCTGCAAGTCGTGTTG
TaSUT2B   TAGGAGCACCTCTGTCGGTTACGTACAGTATACCATATGCGATGGCTGCAAGTCGTGTTG
          ************************************************************

AesSUT2   AAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGCATTCTTAATTTATCTATTGTCATAC
TaSUT2B   AAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGCATTCTTAATTTATCTATTGTCATAC
          ************************************************************

AesSUT2   CACAGATCATCGTGTCGCTGGGCAGCGGGCCATGGGACCAGCTCTTCGGCGGAGGGAACG
TaSUT2B   CACAGATCATCGTGTCGCTGGGCAGCGGGCCATGGGACCAGCTCTTCGGCGGAGGGAACG
          ************************************************************

AesSUT2   CGCCATCCTTCTGGGTGGCTGCCGCAGCCTCCTTTGTGGGCGGGCTGGTAGCCATCCTGG
TaSUT2B   CGCCATCCTTCTGGGTGGCTGCCGCAGCCTCCTTTGTGGGCGGGCTGGTAGCCATCCTGG
          ************************************************************

AesSUT2   GGCTCCCGCGGGCCCGGCTCGGGCCGAAGAAGAAAACCACCCAACGATGATGATTCGAGC
TaSUT2B   GGCTCCCGCGGGCCCGGCTCGGGCCGAAGAAGAAAACCACCCAACGATGATGATTCGAGC
          ************************************************************

AesSUT2   GAGTATACAGAGACATTGTTATGTTGCAGACAAGGAAGGAAGGCAGGTTGTACACTGGTT
TaSUT2B   GAGTATACAGAGACATTGTTATGTTGCAGACAAGGAAGGAAGGCAGGTTGTACACTGGTT
          ************************************************************

AesSUT2   CATTCTTTGCACACTGGTTCCAAGACACCGGGAGCGCTGCCTACACTACACTAGGCATGG
TaSUT2B   CATTCTTTGCACACTGGTTCCAAGACACCGGGAGCGCTGCCTACACTACACTAGGCATGG
          ************************************************************

AesSUT2   GGGGTCTGTATATTGTCATCATCAAGTTGGATTGGGTTCAAATTCAAGGCCAAAAGTAGG
TaSUT2B   GGGGTCTGTATATTGTCTTCATCAAGTTGGGTTCGGTTCAAATTCAAGGCCAAAAGTAGG
          ***************** ************ ** *************************

AesSUT2   AGGGCTCTGGAATTTTCCTCCATGGGCCCTAGCGTCGCCCTTTTTCCGCTTGTCGGTGTG
TaSUT2B   AGGGCTCTGGAATTTTCCTCCATGGGCCCTAGCGTCGCCCTTTTTCCGCTTGTCGGTGTG
          ************************************************************

AesSUT2   AAAAATGGTCAGTCATATGTTGTTGTATATT-GGATATTTTGCCATTTTGTTGGAGATTA
TaSUT2B   AAACATGGTCAGTCATATGTTGTTGTATATTCGGATATTTTGTCATTTTGTCGGAGATTA
          *** ************************** ********** ******** ********

AesSUT2   TCAACCAATCAATAAAAGTTGCAGGTTTTTGTTTGAAA
TaSUT2B   TCAATCAATCAATAAAAGTTGCAGGTTTTTGTTTGAAA
          **** ********************************
```

Figure 3.23 Alignment of the nucleotide sequences of *SUT2* identified from *Aegilops speltoides* (B genome donor; *AesSUT2*) and *Triticum aestivum* cv. AC Andrew (*TaSUT2B*).

3.3.4 Phylogeny of TaSUT2

Generation of a phylogenetic tree using Molecular Evolutionary Genetic Analysis (MEGA, version 5) software (http://www.megasoftware.net/) based on amino acid sequence homology among known monocot and dicot SUTs showed that TaSUT2 is

more closely related to SUT proteins in the SUT-4 subfamily, which mainly contains the low affinity dicot SUT4 proteins and the monocot SUT2 proteins such as those derived from barley and rice (Figure 3.24). Phylogentic analysis of the SUT proteins also showed a very clear division of all known SUTs into 5 groups (Braun and Slewinski, 2009)

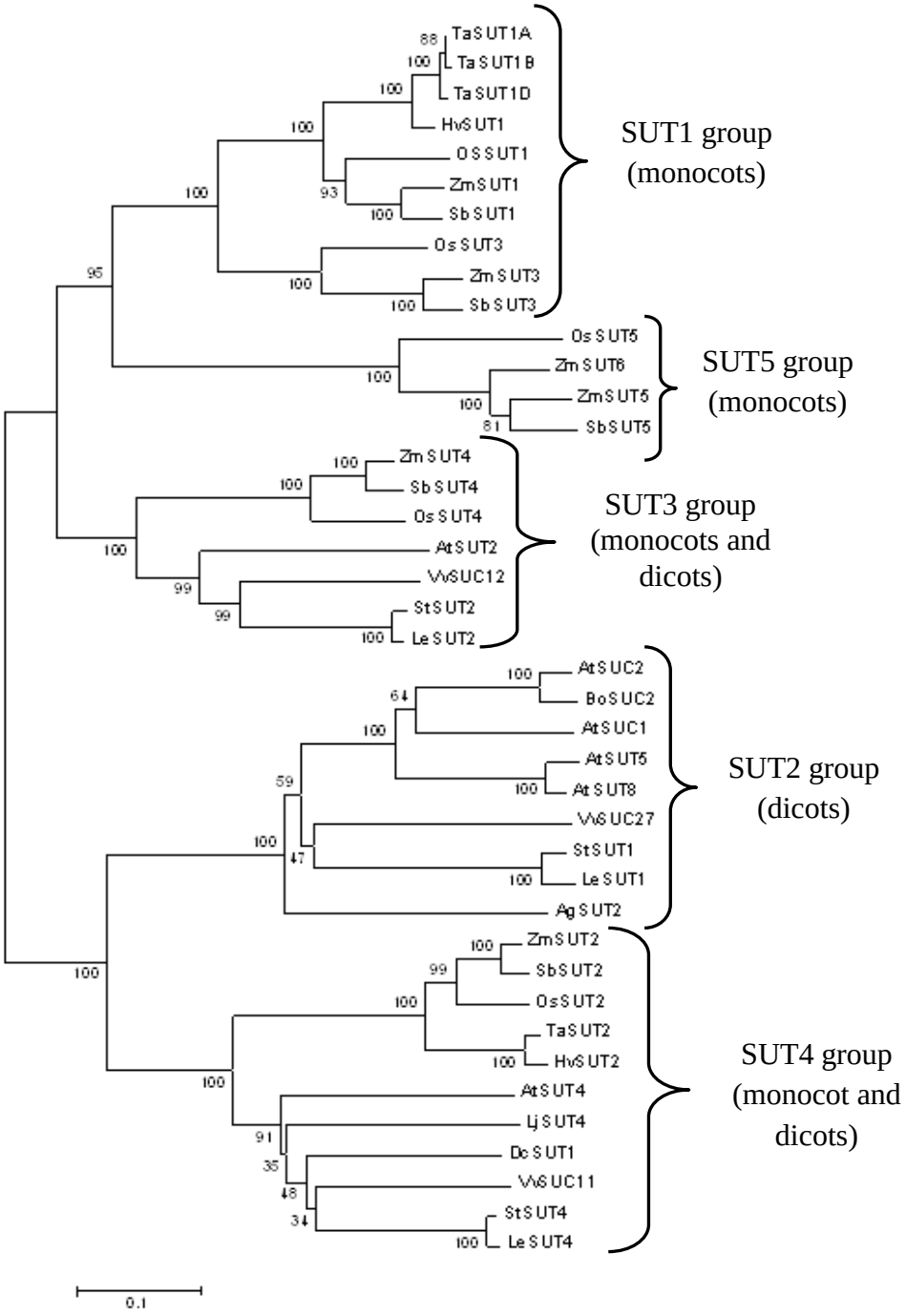


Figure 3.24 Phylogenetic tree based on amino acid sequence similarity of TaSUT2 with other monocotyledonous and dicotyledonous SUTs (generated by using Molecular Evolutionary Genetic Analysis [MEGA, version 5] software). The evolutionary history was inferred using the Neighbor-Joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site. The analysis involved 40 amino acid sequences. The GenBank accession numbers for the aligned peptide sequences are *Apium graveolens*, AgSUT2, AF167415; *Triticum aestivum*, TaSUT1A, AAM13408; TaSUT1B, AAM13409.1; TaSUT1D, AAM13410.1; *Hordeum vulgare*, HvSUT1, CAJ20123.1; HvSUT2, CAB75881.1; *Oryza sativa*, OsSUT1, BAI83443.1; OsSUT2, BAC67163.1; OsSUT3, BAB68368.1; OsSUT4, BAC67164.1; OsSUT5, BAC67165.1; *Zea mays*, ZmSUT1, NP_001104840; ZmSUT2, AAT51689; ZmSUT3, ACF86653.1; ZmSUT4, AATS91375.1; ZmSUT5, ACF85284.1; ZmSUT6, ACF86653.1; *Solanum tuberosum*, StSUT1, CAA48915.1; StSUT2, AAP43631.1; StSUT4, AAG25923.2; *Arabidopsis thaliana*, AtSUC1, CAA53147.1; AtSUC2, CAA53150.1; AtSUT2, AAC32907.1; AtSUT4, AAG09191.1; AtSUT5, BAB11624.1; AtSUT8, AAC69375.1; *Lycopersicon esculentum*, LeSUT1, CAA57726.1; LeSUT2, AAG12987.1; LeSUT4, AAG09270.1; *Vitis vinifera*, VvSUC11, AAF08329.1; VvSUC12, AAF08330.1; VvSUC27, AAF08331.1; *Daucus carota*, DcSUT1A, CAA76367.1; *Sorghum bicolor*, SbSUT1, ACY69230.1; SbSUT2, XX_00243677.1; SbSUT3, XP_002467275.1; SbSUT4, EES06059.1; SbSUT5, XP_002454058.1; *Brassica oleracea*, BoSUC2, AAL58072.1; *Lotus japonica*, LjSUT4, AJ538041.

3.3.5 Recombinant protein expression

Using a protein molecular weight calculator (http://www.sciencegateway.org/tools/proteinmw.htm), the TaSUT2 protein was determined to have a molecular weight of 53.87 Kilodalton (kDa). Amplification of the coding sequence of *TaSUT2* using forward and reverse primers linked with *EcoR1* and *Xho1* recognition sites, respectively, produced a DNA fragment of expected size (Figure 3.25). Cloning the PCR product into pENTR1A entry vector followed by transformation of *E.coli* (DH5α) cells with the resulting construct produced colonies. Screening of the colonies using restriction site linked primers designed from the 5' and 3' ends of putative *TaSUT2* gene identified positive colonies that contained the inserted putative full length

cDNA of *TaSUT2* (Figure 3.26A). Furthermore, digestion of the pENTR1A plasmid containing the insert with *EcoR1* and *Xho1* yielded two fragments, one corresponding to the *TaSUT2* insert (1518 bp) and the other one to the pENTR1A vector (3.8 Kb; Figure 3.26B).

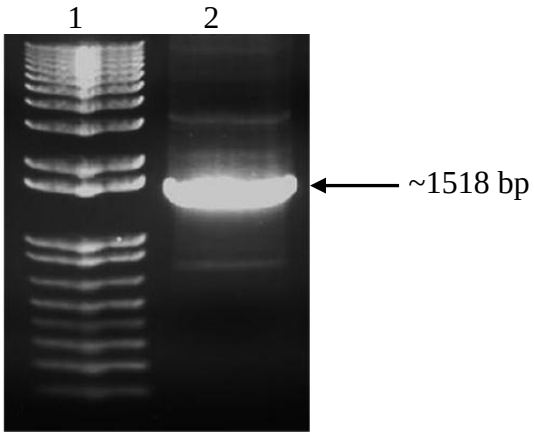


Figure 3.25 PCR product of the coding sequence of *TaSUT2* amplified from cv. AC Andrew using *EcoRI* and *XhoI* restriction site linked forward and reverse primers, respectively. The PCR product was cloned into pENTR1A.

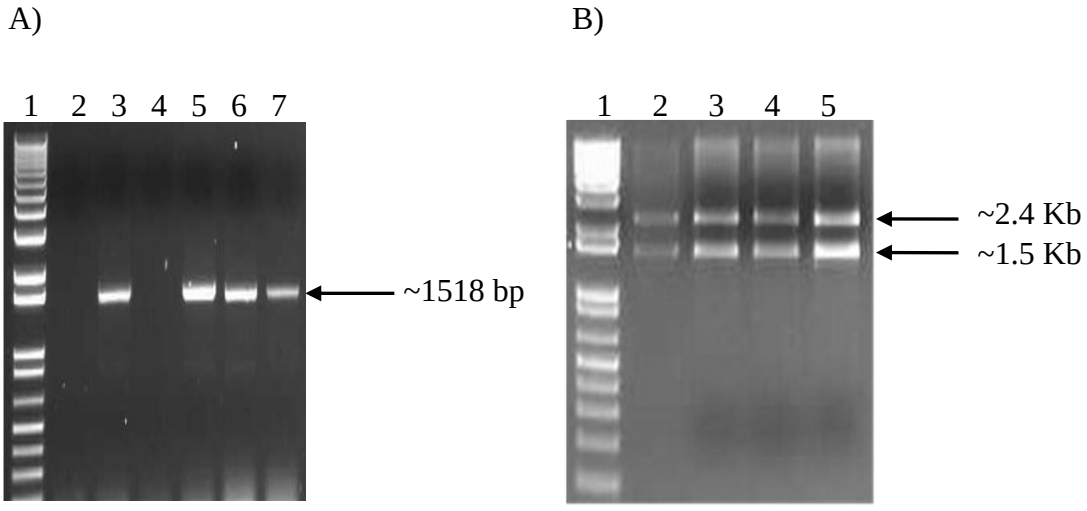


Figure 3.26 PCR screening of *E.coli* (DH5α) cells transformed with the coding cDNA fragment of *TaSUT2*-pENTR1A construct (A); lanes 3, 5, 6 and 7 show DNA fragments of expected size (1518 bp) amplified from the positive colonies. DNA fragments corresponding to (1.5 Kb) and pENTR1A (2.4 Kb) produced by digestion of the *TaSUT2*-pENTR1A construct with *EcoR1* and *Xho1* (B; lanes 2-5). Lane 1 (A, B) corresponds to DNA marker (1Kb+).

Transferring the *TaSUT2* fragment from pENTR1A into a destination (expression) vector, pDEST14, and subsequent transformation of *E.coli* (expression strain BL21) cells with the resulting *TaSUT2*-pDEST14 construct produced expression colonies. PCR screening of the expression colonies with primers designed from the 5' and 3' ends of the putative *TaSUT2* gene yielded an expected band size of 1518 bp (Figure 3.27), verifying the successful insertion of *TaSUT2* into pDEST14. Extraction of the crude proteins from the cells expressing TaSUT2 and subsequent analysis with SDS-PAGE yielded a protein with expected molecular weight of ~54 kDa (Figure 3.28).

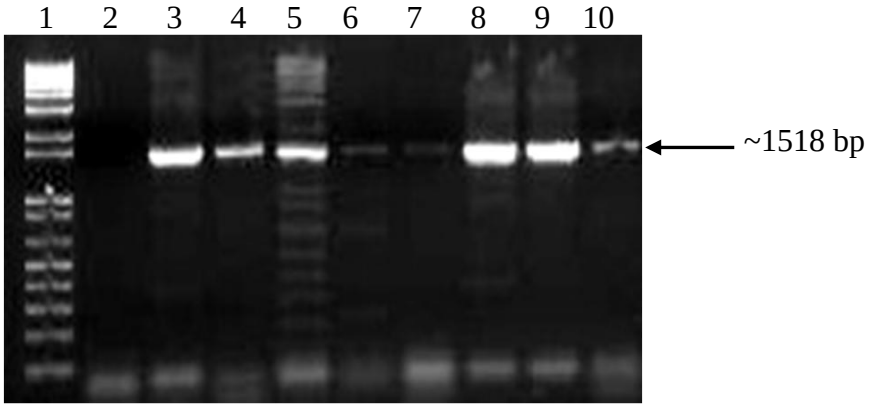


Figure 3.27 PCR screening of *E.coli* (BL21) cells transformed with TaSUT2-pDEST14; lanes 3, 4, 5, 8, 9, 10 show fragments of expected size (1518 bp) amplified from positive colonies. Lane 1 is for DNA marker to determine PCR product size.

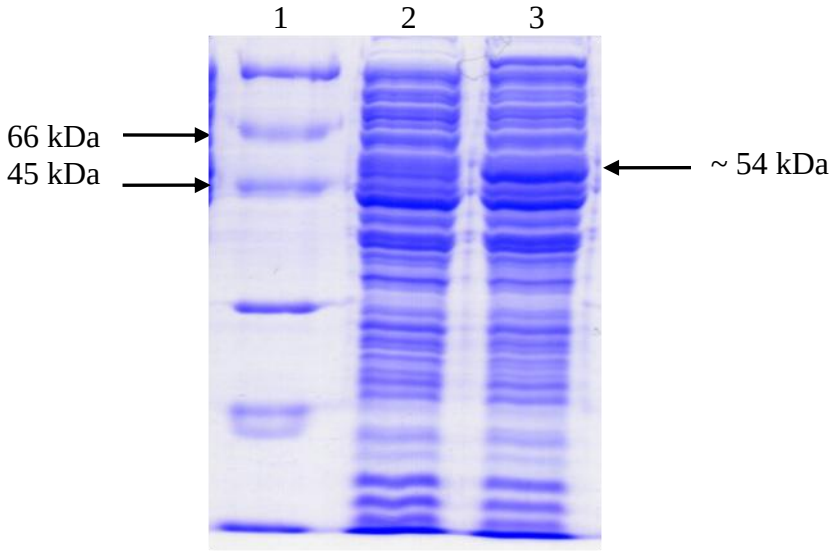


Figure 3.28 SDS-PAGE analysis of TaSUT2 from crude protein samples not induced (lane 2) and induced (lane 3) with IPTG. Lane 1 corresponds to protein marker.

3.3.6 Functional expression in yeast

The functionality of *TaSUT2* was determined by transforming the yeast strain SUSY7/ura3 cells with expression vector (pDR196) containing *TaSUT2*. Amplification of the ORF of *TaSUT2* using restriction site linked forward (*Spe1*-TaSUT2-F) and reverse (TaSUT2-R-*Xho1*) primers produced a DNA fragment of expected size (Figure 3.29). Digestion of both the amplified fragment and the pDR196 vector with *Spe1* and *Xho1* enzymes yielded complementary sites required to ligate the two components (Figure 3.30A). Transformation of the ligation product (TaSUT2-pDR196) with *E.coli* (DH5α) cells produced colonies. Positive expression colonies containing the *TaSUT2*-pDR196 construct were identified by PCR using the restriction site linked primers (Figure 3.30B). Nucleotide sequencing of the inserted *TaSUT2* cDNA fragments

following plasmid isolation confirmed that the inserted sequence has 100% similarity with the coding sequence of *TaSUT2*, except that it contained an additional 14 bp due to the restriction sites linked to it.

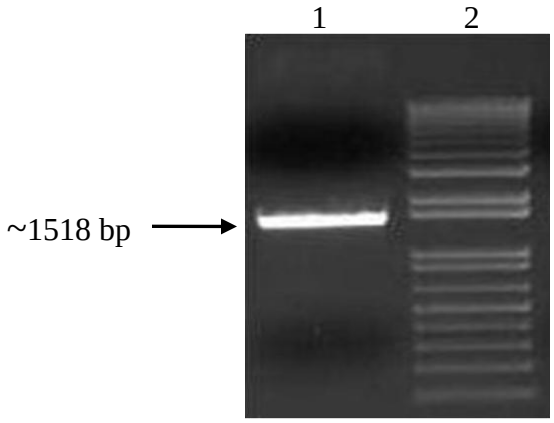


Figure 3.29 PCR product of *TaSUT2* ORF amplified from cv. AC Andrew using *Spe1* and *Xho1* restriction site linked forward and reverse primers (lane 1), respectively. Lane 2 corresponds to DNA marker (1Kb+).

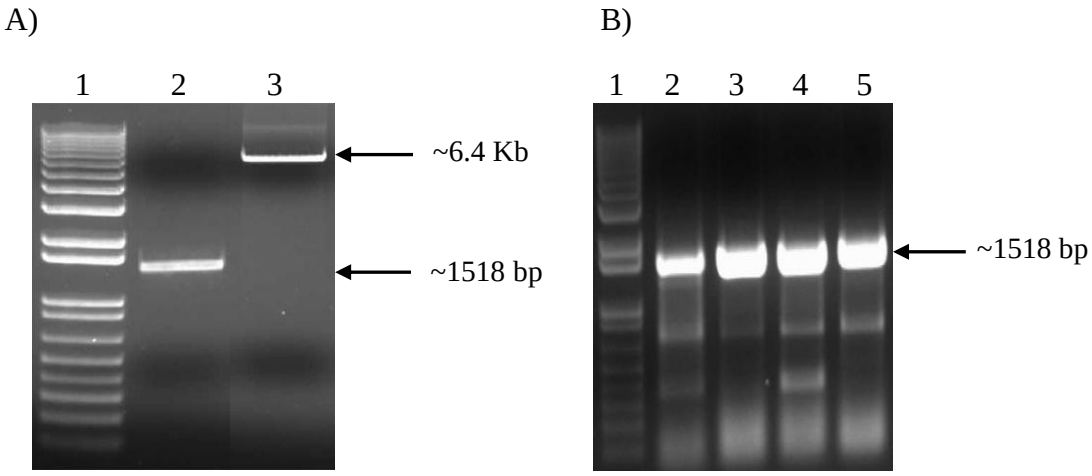


Figure 3.30 DNA fragments corresponding to *TaSUT2* (1.5 Kb, lane 2) and pDR196 (6.4 Kb; lane 4) after digestion with *Spe1* and *Xho1* enzymes. PCR screening of *E.coli* (DH5α) cells transformed with TaSUT2-pDR196 construct (B). Lanes 2, 3, 4 and 5 show DNA fragments of expected size (1518 bp) amplified from the positive colonies. Lane 1 (A, B) corresponds to DNA marker (1Kb+).

Transformation of yeast strain SUSY7/ura3 cells with TaSUT2-pDR196 construct enabled the cells to uptake sucrose and grow on a medium containing sucrose as a sole source of carbon (Figure 3.31). Screening of the SUSY7/ura3 cells grown on sucrose containing media by PCR using the primers described above yielded DNA fragment with expected size (Figure 3.32), verifying the successful transformation of the cells with TaSUT2-pDR196 construct. Moreover, cells transformed with pDR196 containing the potato *SUT1* gene, used as a positive control in this study, successfully grew on sucrose containing medium (Figure 3.31). The SUSY7/ura3 cells transformed with the empty pDR196 did not grow on sucrose containing medium (Figure 3.31). In contrast, the yeast cells of SUSY7/ura3 transformed with either the empty pDR196 or the one containing the *SUT* genes exhibited successful growth on glucose containing medium (Figure 3.31).

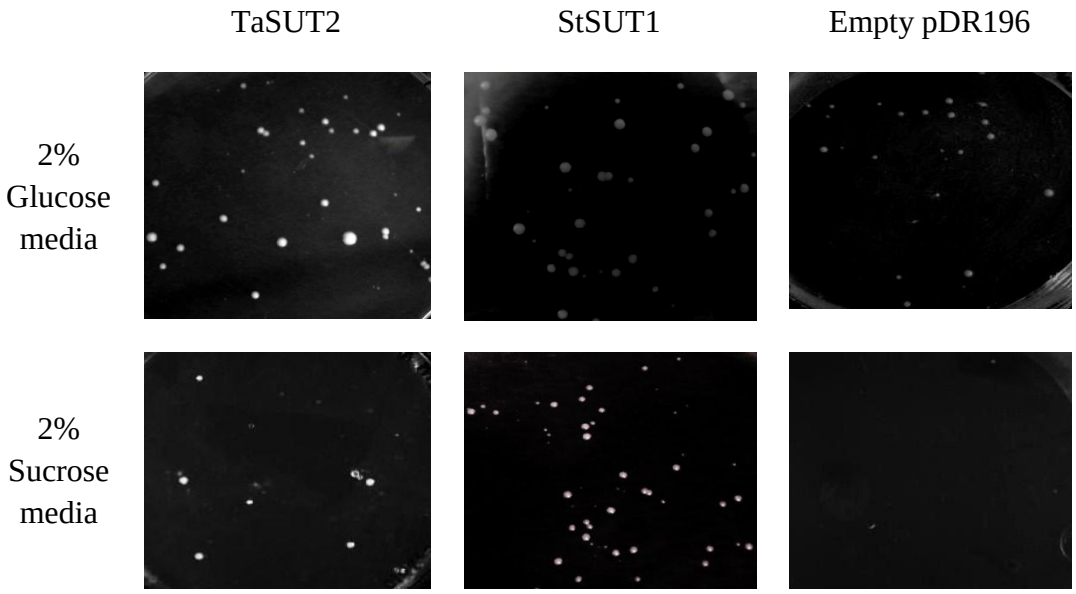


Figure 3.31 Growth of SUSY7/ura3 yeast cells transformed with TaSUT2-pDR196 and StSUT1-pDR196 (positive control) on 2% sucrose and 2% glucose containing medium. Cells transformed with the empty pDR196 vector (negative control) did not show any growth on sucrose containing medium.

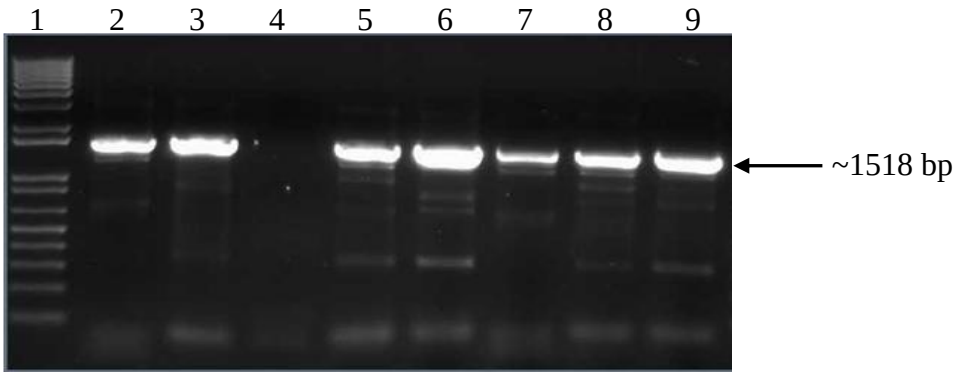


Figure 3.32 PCR screening of SUSY7/ura3 cells transformed with TaSUT2-pDR196 construct; lanes 2, 3, 5, 6, 7, 8 and 9 show DNA fragments of expected size (1518 bp) amplified from the positive colonies. Lane 1 corresponds to DNA marker (1Kb+).

3.3.6.1 Functionality of three homeologues of TaSUT2

Transformation of the yeast cells of SUSY7/ura3 strain with expression vector (pDR196) containing the full length coding sequences of each of *TaSUT2* homeologues identified from cv. AC Andrew (*TaSUT2A, TaSUT2B* and *TaSUT2D*) enabled the cells to uptake sucrose and grow on a medium containing sucrose as a sole carbon source (Figure 3.33).

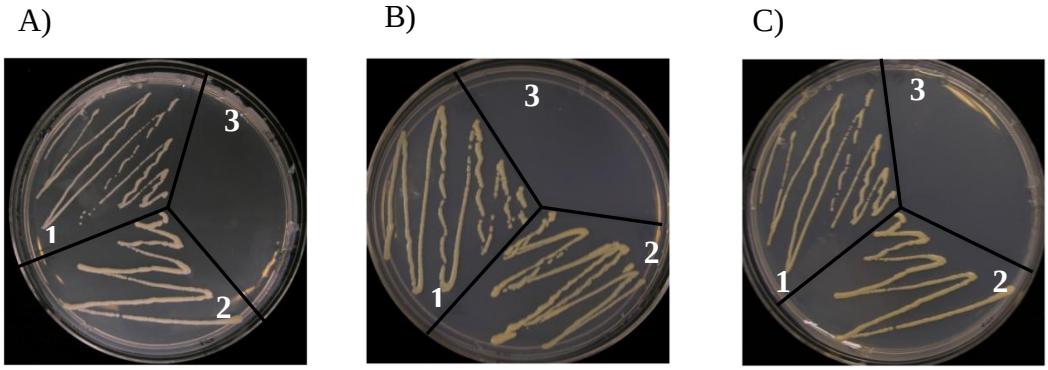


Figure 3.33 Growth of SUSY7/ura3 yeast cells transformed with TaSUT2A-pDR196 (A1) TaSUT2B-pDR196 (B1), TaSUT2D-pDR196 (C1) and the positive control StSUT1-pDR196 (A2, B2 and C2). No growth of SUSY7/ura3 yeast cells was observed with those transformed with the empty pDR196 (negative control; A3, B3 and C3).

3.3.7 Expression analysis of *TaSUT2* in source and sink tissues

The spatio-temporal expression patterns of *TaSUT2* were investigated in developing grains, non-foliar vegetative tissues including the peduncle, glumes, lemma and palea (5, 10, 15, 20, 25 and 30 DAA), and foliar vegetative tissues including the flag leaf blade and flag leaf sheath (at heading [-8 DAA] and 5, 10, 15, 20, 25 and 30 DAA) by using real time qPCR. Amplification of cDNA samples from cv. AC Andrew using *TaSUT2* and *TaActin* specific forward and reverse primers produced a single product with expected length of 130 bp and 137 bp, respectively (Figure 3.34). The transcripts of *TaSUT2* were detected in all tissues examined in this study.

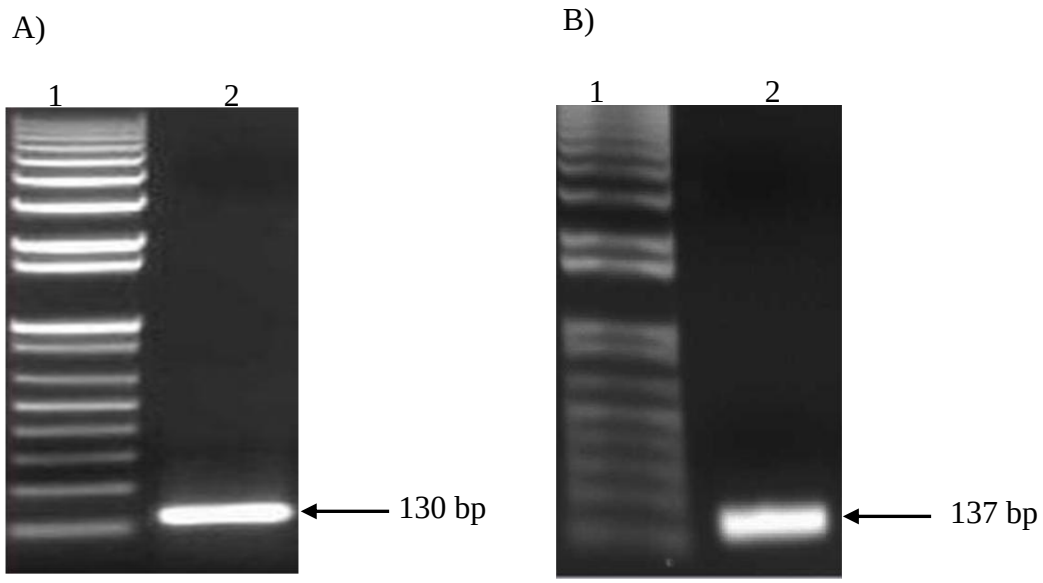
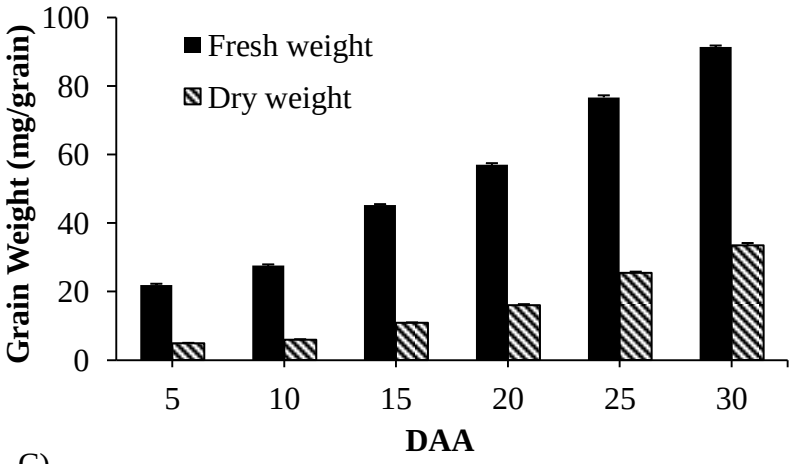


Figure 3.34 PCR products of *TaSUT2* (A; lane 2) and *TaActin* (B; lane 2) amplicons produced by amplification of cDNA samples with *TaSUT2* and *TaActin* specific primers. Lane 1 (A, B) corresponds to DNA marker (1Kb+).

A)

B)



C)

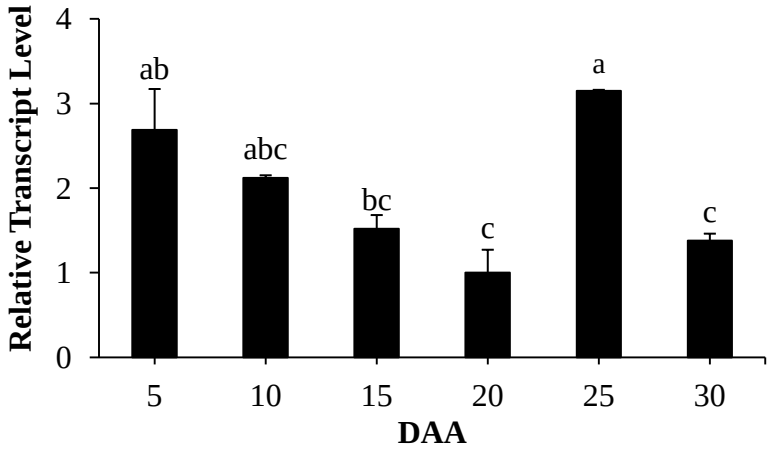


Figure 3.35 Developing grains of wheat cv. AC Andrew from 5 to 30 DAA (A) with their fresh and dry weights (B). Data are means ± SE, n=20. Relative abundance of *TaSUT2* transcript in developing wheat grains (5, 10, 15, 20, 25 and 30 DAA; C). The transcript levels were compared across developmental stages and tissues using the average transcript level in 20 DAA grain samples, which was set to 1, as calibrator. Data are means ± SE, n = 2 to 3. Different letters indicate significant difference among developmental stages using Tukey's LSD, $p<0.05$.

Grain fresh and dry weights increased almost linearly as the grains develop from 5 to 30 DAA (Figure 3.35). Initiation of fresh and dry matter accumulation by 5 DAA was associated with relatively higher transcript abundance of *TaSUT2*. Further increase in fresh and dry matter accumulation as the grain develops through 20 DAA was, however, accompanied by a gradual decline in the abundance of *TaSUT2* transcripts. As grain growth continued to increase from 20 to 25 DAA, the level of *TaSUT2* transcripts exhibited a 3-fold increase, attaining a similar level observed in 5 DAA grains. Further grain growth through 30 DAA was associated with a 1.5-fold decline in the transcript abundance of *TaSUT2* (Figure 3.35D). Overall, the expression of *TaSUT2* in grains was significantly higher ($P<0.05$) at earlier and later stages than mid-stage of development.The expression of *TaSUT2* in the flag leaf blade was significantly higher at heading and during the early periods of grain development (5 to 10 DAA; $P<0.05$) than the later stages, when it was maintained at a lower expression through 30 DAA (Figure 3.36). However, its expression in the flag leaf sheath was almost similar during grain development, though higher expression ($P<0.05$) was evident at 10 and 25 DAA (Figure 3.36). In the peduncle, the expression of *TaSUT2* was maintained at similar level at all stages examined (Figure 3.37).

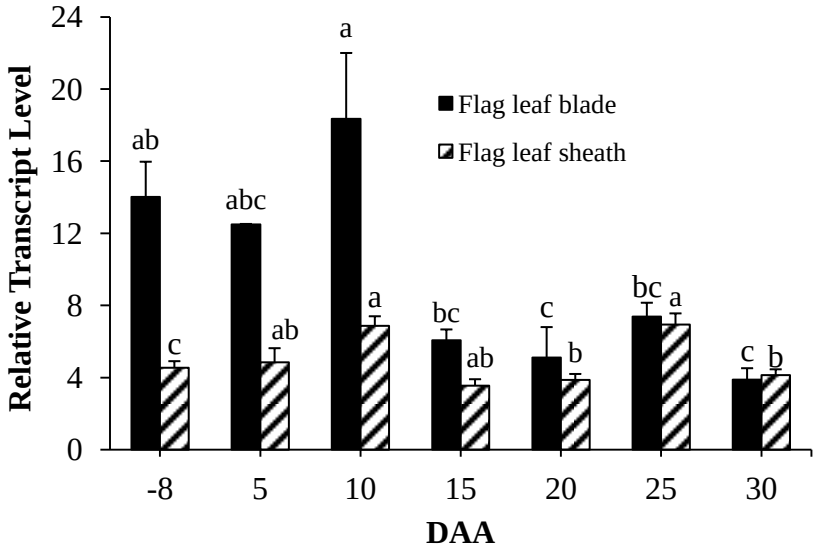


Figure 3.36 Relative abundance of *TaSUT2* transcript in the flag leaf blade and sheath during grain development in wheat (-8, 5, 10, 15, 20, 25 and 30 DAA). The transcript levels were compared across developmental stages using the average transcript level in 20 DAA grain samples, which was set to 1, as calibrator. Data are means ± SE, n = 2 to 3. Different letters within tissues indicate significant difference among developmental stages using Tukey's LSD, $p<0.05$.

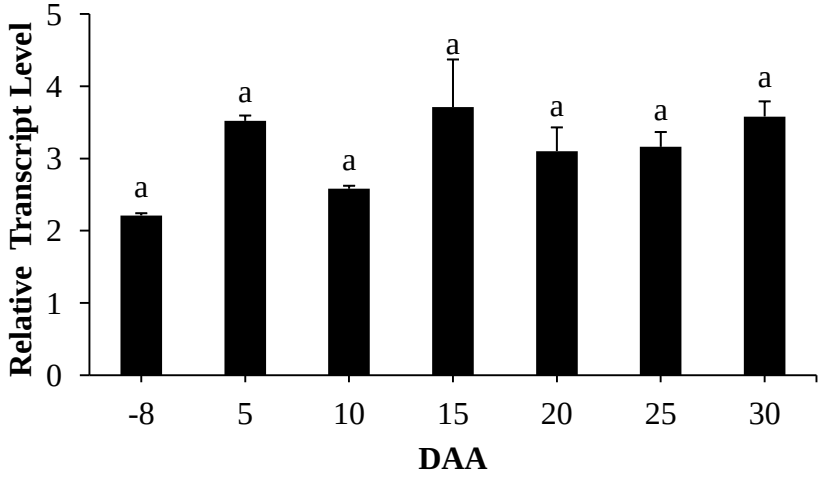


Figure 3.37 Relative abundance of *TaSUT2* transcript in the peduncle during grain development in wheat (-8, 5, 10, 15, 20, 25 and 30 DAA). The transcript levels were compared across developmental stages and tissues using the average transcript level in 20 DAA grain samples, which was set to 1, as calibrator. Data are means ± SE, n = 2 to 3. Different letters indicate significant difference among developmental stages using Tukey's LSD, $p<0.05$.

TaSUT2 exhibited similar expression level and pattern between the non-foliar lemma and palea tissues during grain development (Figure 3.38). Its expression in the glume also followed a similar pattern, but with an expression level slightly lower than that observed in the lemma and palea tissues (3-fold; Figure 3.38). There was no significant difference ($P<0.05$) in *TaSUT2* expression in each of the glume, lemma and glume tissues across the developmental stages examined in this study.

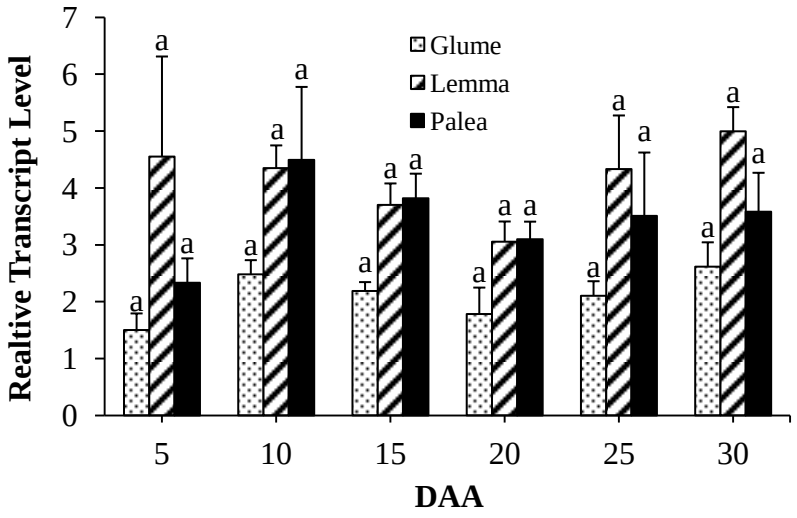


Figure 3.38 Relative abundance of *TaSUT2* transcript in the glume, lemma and palea during grain development in wheat (5, 10, 15, 20, 25 and 30 DAA). The transcript levels were compared across developmental stages and tissues using the average transcript level in 20 DAA grain samples, which was set to 1, as calibrator. Data are means ± SE, n = 2 to 3. Different letters within tissues indicate significant difference among developmental stages using Tukey's LSD, $p<0.05$.

3.4 Discussion

The biosynthesis of starch, which is the major determinant of yield in cereal crops, is dependent on the translocation of sucrose, the major form of transported sugar from the source to the sink organs in plants. Sucrose active transport across the cellular membrane barriers is mediated by sucrose transporter (SUT) proteins. Genes encoding SUT proteins have been identified and characterized from a number of monocot crop species including rice, maize and barley. In hexaploid wheat, however, only one *SUT* (*TaSUT1*) has thus far been identified and functionally characterized (Aoki et al., 2002, 2004, 2006). This thesis reports the identification and functional characterization of cDNAs of a new SUT (*SUT2*) from the three genomes of hexaploid wheat cv. AC Andrew (*TaSUT2A*, *TaSUT2B* and *TaSUT2D*). The *TaSUT2A* and *TaSUT2B* genes have ORF of 1518 bp, whereas *TaSUT2D* encodes an ORF of 1524 bp encoding proteins with 506 and 508 amino acids, respectively, and molecular mass of approximately 54 kDA. Similar to that observed in the three homeologues of *TaSUT1* (Aoki et al., 2002), the *TaSUT2A*, *TaSUT2B* and *TaSUT2D* genes are found to be different mainly in the 3' untranslated region. BLAST search of GenBank with the nucleotide sequences of the three homeologues of *TaSUT2* indicated that they have high sequence identity with the genes of other known cereal *SUT*s including that of barley (*HvSUT2*; 92%), rice (*OsSUT2*; 80%) and maize (*ZmSUT2*; 80%). Their predicted amino acid sequences have high similarity with other cereal SUTs including those derived from barley (HvSUT2; 97%), rice (OsSUT2; 85%) and maize (ZmSUT2; 83%). Analysis of the predicted amino acid sequences of the three *SUT2* genes revealed that their corresponding proteins are members of the GPH cation symporter family to which all the other plant SUTs belong

(Lalonde et al., 2004). Further *in silico* analysis of the topology of the predicted amino acid sequences of TaSUT2A, B and D showed that their corresponding proteins contain the 12 transmembrane helices, coined as a distinct characteristic feature of all known SUTs (Marger and Saier, 1993; Shiratake, 2007), and the consensus sequence derived from the highly conserved region of functional plant SUTs (Aoki et al., 2003; Figure 3.17). Furthermore, the histidine residue, which is considered to be responsible for the binding of sucrose during its transport and shown to be conserved in all known plant SUTs (Lu and Bush, 1998; Aoki et al., 2003), is found in the first extracellular loop of TaSUT2 proteins (His-61 in case of TaSUT2A and TaSUT2B, and His-63 of TaSUT2D). These results overall suggest that the newly isolated *SUT2* genes encode sucrose transporters.

Generation of a phylogenetic tree on the basis of amino acid sequences of known SUTs from both dicot and monocot species showed the clustering of all the SUTs into five groups (Braun and Slewinski, 2009; Kuhn, 2011). The TaSUT2 did not cluster into groups that exclusively contain only SUTs derived from monocot, SUT1 and SUT5 groups; rather it fall into the SUT4 group that contains both dicot and monocot SUTs including HvSUT2, OsSUT2, SbSUT4 and ZmSUT2 (Braun and Slewinski, 2009; Aoki et al., 2012). The peptide sequences of TaSUT2A, TaSUT2B and TaSUT2D showed only 44.28%, 44.09 and 44.01% similarity with that of TaSUT1A, TaSUT1B and TaSUT1D that are classified in SUT1 group, respectively, leading to a phylogenetically distant relationship and their assignment into different SUT groups. Likewise, HvSUT1, which is classified in SUT1 group, and HvSUT2 share only 42.2% similarity in their amino acid sequence (Weschke et al., 2000). The SUTs are believed to be localized mainly on the

plasma membrane of sieve elements and companion cells (Truernit, 2001; Reinders et al., 2002). However, a recent proteomic study has shown that both HvSUT2 and AtSUT4, belong to SUT4 group, are expressed in vacuolar membrane (Endler et al., 2006). Furthermore, LjSUT4 *(Lotus japonicas)* is found to be a vacuolar SUT (Reinders et al., 2008). It is therefore likely that TaSUT2A, TaSUT2B and TaSUT2D are vacuolar SUTs that are involved in sucrose efflux from the vacuole across the tonoplast (Endler et al., 2006; Kuhn, 2011). However, this needs to be verified with subcellular localization study.

Though the sequence homology with other known functional SUTs, predicted protein topology and presence of conserved domains/motifs provided an insight into the function of TaSUT2A, TaSUT2B and TaSUT2D as SUTs, to determine their actual function we performed a complementation study via heterologous expression of TaSUT2A, TaSUT2B and TaSUT2D in SUSY7/*ura3* strain of yeast (*Saccharomyces cerevisiae*), which has been used as a tool of choice to verify the biological functionalities of several plant SUTs. Since this strain of yeast is a mutant that cannot utilize external sucrose, their ability to uptake sucrose and grow on medium containing sucrose as a sole carbon source following transformation with *TaSUT2A*, *TaSUT2B* and *TaSUT2D* confirmed that their corresponding proteins are functional SUTs. In order to gain further insight into the physiological roles of *TaSUT2*, we analyzed its expression in both source (flag leaf blade and sheath, peduncle, glume, lemma and plaea) and sink (developing grains) organs immediately before and during reproductive development. The *TaSUT2* gene is found to be expressed nearly at equal level in all tissues analyzed, suggesting its multifunctional role including phloem loading and unloading, and retrieval

of sucrose diffused to the vascular tissue during its transport. Likewise, the *SUT2* of rice is expressed in both source and sink tissues, though its expression in the seeds was limited to their early stage (1-7 DAA) of development (Aoki et al., 2003). The expression of *TaSUT2* in seeds was relatively higher during the early (5-10 DAA) and late (25 DAA) stages of their development, whereas *TaSUT1* was highly expressed during the rapid phase of grain filling (16-20 DAH; Aoki et al., 2002), when *TaSUT2* expression is lower. Localization of SUTs of barley (*HvSUT1* and *HvSUT2)* and rice (*OsSUT1*) in the nucellar projection and transfer cells of developing endosperms (Weschke at al., 2000; Takeda et al., 2001) suggests that SUT proteins of cereals are primarily involved in importing sucrose from grain apoplasm across the aleurone layer into the developing endosperm, where starch is synthesized. It is therefore likely that TaSUT1 and TaSUT2 play temporally distinct but functionally overlapping roles during grain filling. The decrease in *TaSUT2* expression after 25 DAA can be attributed to less sink demand during later stages of grain filling.

Current research on SUT proteins in cereals is focused mainly on their role in importing sucrose to the developing grain due to the economic importance of this organ. However, the transcripts of *SUT*s are also found in the vegetative tissues of a number of cereal crops including rice (Hirose et al., 1997; Aoki et al., 2003), barley (Weschke et al., 2000), wheat (Aoki et al., 2002, 2004, 2006) and maize (Aoki et al., 1999). During reproductive development, grain filling in wheat is supported mainly by photoassimilates produced in the flag leaf (Ali et al., 2010). Thus, the relatively higher expression of *TaSUT2* in the source flag leaf blade especially during the early stage (-8 to 10 DAA) of grain development as compared to the other tissues analyzed in this study may suggest

the role of TaSUT2 in the active phloem loading of sucrose into the sieve element complex of phloem as the mesophyll and vascular cells of wheat possess no symplastic connections (Aoki et al., 2004). The higher expression of *TaSUT1* in the flag leaf blade at 4 days before heading and its decline by 12 days after heading where heading starts before 8 - 10 days prior to anthesis (Aoki et al., 2002) suggests the complimentarity of these two genes in phloem loading. Likewise, higher expression of *OsSUT2* was observed in the source flag leaf blade of rice as compared to flag leaf sheath and internodes (Sun et al., 2008). Furthermore, rice *SUT1* exhibit relatively low level of expression in source leaves and antisense repression of *OsSUT1* did not have any effect on phloem loading of sucrose (Scofield et al., 2002). The decline in the transcript abundance of *TaSUT2* in the flag leaf blade during the later stages of grain filling can be attributed to the senescence of leaves (Schynder, 1993; Kong et al., 2010).

The leaf sheaths of mature wheat plant act as a reservoir for temporary storage of excess carbohydrates in the form of water soluble sugars such as fructans, and resupply carbon to the developing grain when photoassimilate production is decreased due to leaf senescence or increased sink demand (Schnyder, 1993). Therefore, the presence of *TaSUT2* transcripts in the flag leaf sheath during grain development suggests its role in remobilizing photoassimilates to the developing grains. Consistent with this, approximately 50% of photoassimilates contributing to grain filling are remobilized out of tissues used as temporary storage (Schnyder, 1993; Scofield et al., 2007). The peduncle of wheat, which is the uppermost internode below the spike, serves as a pathway for sucrose transport from the leaf to developing grain (Kong et al., 2010). The detection of similar levels of *TaSUT2* expression in peduncle during grain development

suggests its role in retrieving sucrose that leaked out of phloem apoplasm during transportation, similar to the function assigned to *TaSUT1* present in peduncle (Truernit, 2001; Aoki et al., 2006). As the peduncle and leaf sheath tissues of wheat are partly exposed to the incoming radiation and produce photoassimilates that ultimately contributes approximately 9-12% of wheat grain dry matter (Wang et al., 2001; Kong et al., 2010), it is likely that TaSUT2 is also involved in phloem loading of sucrose produced in these tissues for its transport to the developing grains. It has been reported previously that photosynthesis in the non-foliar vegetative tissues that comprise the wheat spike (glume, lemma and palea) contributes 10-44% of photoassimilates destined to wheat grain (Kriedemann, 1966; Wang et al., 2001). The maintenance of almost similar abundance of *TaSUT2* transcripts in these tissues. during grain development suggests the role of TaSUT2 in the active phloem loading of sucrose produced in these tissues.

In summary, our study demonstrated that *TaSUT2* is a novel gene encoding a true sucrose transporter protein, and analysis of its spatio-temporal expression suggests its role in phloem loading and unloading, and sucrose retrieval processes. However, elucidating its definitive physiological role requires further studies including examining its subcellular localization in both vegetative tissues and developing grains, and generation of mutants and subsequent analysis of the resulting phenotypes.

4.0 GENERAL DISCUSSION AND CONCLUSIONS

Increased diversified use of starch deposited in wheat grains, for example, for bioethanol production, is leading to a growing demand for this crop. Starch yield in wheat grains is determined partly by the amount and efficiency of sucrose transport from the source tissues, where it is synthesized, to the endosperm, where it is used as substrate for starch synthesis. Genes encoding sucrose transporters (SUTs) involved in phloem loading and unloading of sucrose, and retrieval of the sucrose diffused in vascular tissues during transport have been isolated from different crop species including cereal crops. Functional analysis of these genes has advanced our knowledge of sucrose transport in plants. Furthermore, manipulation of these genes has led to enhanced starch yield in a number of crops such as potato and pea. However, only one *SUT (TaSUT1)* has been reported in wheat thus far. Given that SUTs in cereal crop species are encoded by small gene families, it is more than likely that the wheat genome contains additional SUTs. Thus, identifying more wheat SUTs and characterizing their functionality will have a significant contribution in enhancing our understanding of carbon partitioning in wheat, thereby designing tools for improved wheat starch yield.

Using a RACE-PCR approach, this study identified the three homeologues of a new *SUT* gene *(TaSUT2)* from common wheat. The full length sequences of TaSUT2A, TaSUT2B and TaSUT2D are 2036 bp, 2018 bp and 2042 bp with open reading frames of 1518 bp, 1518 bp and 1524 bp encoding 506 amino acids for TaSUT2A and TaSUT2B, and 508 amino acids for TaSUT2D. Searching GenBank with their nucleotide and amino acid sequences showed that the three homeologues have high similarity with the cDNA and protein sequences derived from other known plant SUTs. Further *in silico* analysis of

their amino acid sequence indicated that the TaSUT2s are new member of GPH cation symporter family to which all other known SUTs belong. Phylogenetic analysis of TaSUT2 revealed that it is closely related to previously identified monocot SUTs, HvSUT2 and OsSUT2, which form a separate group of SUTs with SUT4s of dicots species such as that of potato (StSUT4) and Arabidopsis (AtSUT4). The growth of yeast cells of the SUSY7/ura3 strain, which cannot uptake sucrose, transformed with a vector carrying the *TaSUT2A*, *TaSUT2B* or *TaSUT2D* genes on a medium containing sucrose as a sole carbon source revealed the functionality of TaSUT2A, TaSUT2B and TaSUT2D as true SUT proteins. Though our phylogenetic analysis indicated that TaSUT2s belong to SUT4 subfamily that is characterized by low affinity for sucrose, further analysis is required to precisely determine its kinetic properties.

The spatio-temporal expression analysis of *TaSUT2* revealed that its transcripts are detected in all source and sink tissues examined in this study. However, differential levels of expression of this gene were evident across tissues and developmental stages. Unlike *TaSUT1*, which is highly expressed during the rapid phase of grain filling, the expression of *TaSUT2* is relatively higher during the early and later stages of seed development. These results suggest that both *TaSUT1* and *TaSUT2* contribute to sucrose transport at different time periods. The presence of *TaSUT2* transcripts in the flag leaf blade and other non-foliar but photosynthetic tissues such as palea, glume and lemma, which are known to contribute photoassimilates to developing grains, indicates the involvement of TaSUT2 in phloem loading of sucrose. The expression of *TaSUT2* in the peduncle, which acts as a pathway for sucrose transport, suggests its role in retrieval of sucrose leaked out to the phloem apoplasm during its transport. The expression of

TaSUT2 in the leaf sheath, which acts as temporary storage for excess photoassimilates, suggests its role in remobilizing storage reserves. Based on their localization to vacuolar membranes, members of the SUT4 subfamily are suggested to be responsible for sucrose efflux from the vacuole. However, further analysis of *TaSUT2* subcellular localization in both vegetative tissues and developing grains of wheat is required to shed light on its cellular/tissue specific roles. Although the spatio-temporal analysis of *TaSUT2* presented in this study provided an insight into its role, elucidation of the definitive physiological function of this gene requires generation of mutants such as those over- and under-expressing *TaSUT2* and analysis of the resulting mutants in terms of starch accumulation in the grains and other related yield parameters. In conclusion, the findings of this thesis project enhances our understanding of sucrose transport in wheat, and further analysis of the newly identified gene has the potential of providing additional tools to develop high starch yielding wheat cultivars suited for bioethanol production.

LITERATURE CITED

Ajanovic A (2010) Biofuels versus food production: Does biofuels production increase food prices? Energy **36:** 2070-2076

Aldape MJ, Elmer AM, Chao WS, Grimes HD (2003) Identification and characterization of a sucrose transporter isolated from the developing cotyledons of soybean**.** Arch Biochem Biophys **409**: 243-250

Ali MA, Hussain M, Khan MI, Ali Z, Zulkiffal M, Anwar J, Sabir W, Zeeshan M (2010) Source - sink relationship between photosynthetic organs and grain yield attributes during grain filling stage in spring wheat (*Triticum aestivum*). Int J Agric Biol **12:** 509-515

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Blast local alignment search tool. J Mol Biol **215:** 403-10

Aoki N, Scofield GN, Wang XD, Patrick JW, Offler CE, Furbank RT (2004) Expression and localization analysis of the wheat sucrose transporter *TaSUT1* in vegetative tissues. Planta **219:** 176-184

Aoki N, Hirose T, Furbank RT (2012) Sucrose transport in higher plants: from source to sink. In JJ Eaton Rye, BC Tripathy, TD Sharkey, eds, Photosynthesis: plastid biology, energy conservation and carbon assimilation, advances in photosynthesis and respiration, Vol 34. Springer Dordrecht Heidelberg London, New York, pp 703-729

Aoki N, Hirose T, Scofield GN, Whitfeld PR, Furbank RT (2003) The sucrose transporter gene family in rice. Plant Cell Physiol **44:** 223-232

Aoki N, Hirose T, Takahashi S, Ono K, Ishimaru K, Ohsugi R (1999) Molecular cloning and expression analysis of a gene for a sucrose transporter in maize (*Zea mays* L.). Plant Cell Physiol **40:** 1072-1078

Aoki N, Scofield GN, Wang XD, Offler CE, Patrick JW, Furbank RT (2006) Pathway of sugar transport in germinating wheat seeds. Plant Physiol **141:** 1255-1263

Aoki N, Whitfield P, Hoeren F, Scofield G, Newell K, Patrick J, Offler C, Clarke B, Rahman S, Furbank RT (2002) Three sucrose transporter genes are expressed in the developing grain of hexaploid wheat. Plant Mol Biol **50**: 453-462

Araus JL, Tapia L (1987) Photosynthetic gas exchange characteristics of wheat flag leaf blades and sheaths during grain filling. Plant Physiol **85:** 667-673

Austin A (2009) Embracing the alternatives. http://www.ethanolproducer.com/articles /5586/embracing-the-alternatives (accessed on 4 October, 2011)

Ayre BG (2011) Membrane-transport systems for sucrose in relation to whole - plant carbon partitioning. Mol Plant **4:** 377-394

Balat M, Balat H (2009) Recent trends in global production and utilization of bio-ethanol fuel. Appl Energ **86:** 2273-2282

Barker L, Kuhn C, Weise A, Schulz A, Gebhardt C, Hirner B, Hellmann H, Schulze W, Ward J, Frommer WB (2000) SUT2, a putative sucrose sensor in sieve elements. Plant Cell **12**: 1153-1164

Botha CEJ, Cross RHM (1997) Plasmodesmatal frequency in relation to short-distance transport and phloem loading in leaves of barley (*Hordeum vulgare*). Phloem is not loaded directly from the symplast. Physiol Plant **99:** 355-362

Braun DM, Slewinski TL (2009) Genetic control of carbon partitioning in grasses: roles of sucrose transporters and tie-dyed loci in phloem loading. Plant Physiol **149:** 71-81

Buller AHR (1919) The early history of wheat growing in Manitoba. In Essays on wheat, Macmillan company, New York, pp 1-33, http://www.archive.org/details/essays onwheatinc01bull (accessed on 4 October 2011)

Burkle L, Hibberd JM, Quick WP, Kuhn C, Hirner B, Frommer WB (1998) The H1-sucrose co-transporter *NtSUT1* is essential for sugar export from tobacco leaves. Plant Physiol **118:** 59-68

Canada Wheat Board (2011) Western Canada growing season in review - 2011. http://www.cwb.ca/public/en/farmers/grain/crop/ (accessed on 13 October 2011)

Canadian Renewable Fuel Association [CRFA] (2010) Growing beyond oil: delivering our energy future. A report card on Canada's Renewable Fuel Association, http://www.greenfuels.org/uploads/documents/crfareportcardenglish2010final.pdf (accessed on 13 October 2011)

Carpaneto A, Geiger D, Bamberg E, Sauer N, Fromm J, Hedrich R (2005) Phloem-localized, proton-coupled sucrose carrier *ZmSUT1* mediates sucrose efflux under the control of the sucrose gradient and the proton motive force. J Biol Chem **280:** 21437-21443

Carroll A, Somerville C (2009) Cellulosic biofuels. Annu Rev Plant Biol **60:** 165-82

Chandran D, Reinders A, Ward JM (2003) Substrate specificity of the *Arabidopsis thaliana* sucrose transporter AtSUC2. J Biol Chem **278**: 44320-44325

Chang AB, Lin R, Studley WK, Tran CV, Saier MH (2004) Phylogeny as a guide to structure and function of membrane transport proteins. Mol Membr Biol **21:** 171-181

Crowe B (2008) Biofuels: a molecular perspective. http://digitalcommons.unl.edu/ageconundergrad/1 (accessed on 27 September 2011)

Curtis, BC (2002) Wheat in the world. In BC Curtis, S Rajaram, HG Macpherson, eds, Bread wheat improvement and production, Food and Agriculture Organization of the United Nations, Rome http://www.fao.org/docrep /006/y4011e/y4011e04.htm#bm04 (accessed on 3 October, 2011)

Demirbas A (2005) Bioethanol from cellulosic materials: A renewable motor fuel from biomass. Energ Source **27:** 327-337

Demirbas A (2007) Progress and recent trends in biofuels. Prog Energ Combust **33**: 1-18

Demirbas A (2009) Biofuels. In Biofuels: securing the planet's future energy needs, Springer-Verlag London Limited, pp 87-101

Dessureault D (2011) Canada biofuels annuals. http://gain.fas.usda.gov/Recent%20GAIN%20Publications/Biofuels%20Annual_Ottawa_Canada_07-05-2011.pdf (accessed on 30 September 2011)

Eckardt NA (2003) The function of SUT2/ SUC3 sucrose transporters: the debate continues. Plant Cell **14:** 1259-1262

Emes MJ, Bowsher CG, Hedley C, Burrell MM, Scrase-Field ESF, Tetlow IJ (2003) Starch synthesis and carbon partitioning in developing endosperm. J Exp Bot **54:** 569-575

Endler A, Meyer S, Schelbert S, Schneider T, Weschke W, Peters SW, Keller F, Baginsky S, Martinoia E, Schmidt UG (2006) Identification of a vacuolar sucrose transporter in barley and arabidopsis mesophyll cells by a tonoplast proteomic approach. Plant Physiol **141:** 196-207

Eom JS, Cho JI, Reindrs A, Lee SW, Yoo Y, Tuan PQ, Choi SB, Bang G, Park YI, Cho MH, Bhoo SH, An G, Hahn TR, Ward Jm, Jeon JS (2011) Impaired function of the tonoplast-localized sucrose transporter in rice, *OsSUT2*, limits the transport of vacuolar reserve sucrose and affects plant growth. Plant Physiol **157:** 109-119

Feldman M (1997) Origin of cultivated wheat. In AP Bonjean, WJ Angus eds, The world wheat book, a history of wheat breeding, Lavoisier Publishing, France, pp 4-56

Fu C, Mielenz JR, Xiao X, Ge Y, Hamilton CY, Rodriguez M, Chen F, Foston M, Ragauskas A, Bouton J, Dixon RA, Wang ZY (2011) Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. Proc Natl Acad Sci **108:** 3803-3808

Furbank RT, Scofield GN, Hirose T, Wang XD, Patrick J, Offler CE (2001) Cellular localization and function of a sucrose transporter *OsSUT1* in developing rice grains. Aust J Plant Physiol **28:** 1187-1196

Gifford RM, Thorne JH, Hitz WD, Giaquinta RT (1984) Crop productivity and photoassimilate partitioning. Science **225:** 801-808

Gill BS, Friebe B (2002) Cytogenetics, pylogeny and evolution of cultivated wheats. In BC Curtis, S Rajaram, HG Macpherson eds, Bread wheat improvement and production, Food and Agriculture Organization of the United Nations, Rome

http://www.fao.org/docrep/006/y4011e/y4011e04.htm#bm04 (accessed on 3 October, 2011)

Gray KA, Zhao L, Emptage M (2006) Bioethanol. Curr Opin Chem Biol **10:** 141-146

Hahn-Hagerdal B, Galbe M, Gorwa-Grauslund MF, Liden G, Zacchi G. (2006) Bioethanol- the fuel of tomorrow from the residues of today. Trends Biotechnol **24:** 549-556

Harrington GN, Franceschi VR, Offler CE, Patrick JW, Tegeder M, Frommer WB, Harper JF, Hitz WD (1997) Cell specific expression of three genes involved in plasma membrane sucrose transport in developing *Vicia faba* seed. Protoplasma **197:** 160-173

Haupt S, Duncan GH, Holzberg S, Oparka KJ (2001) Evidence for symplastic phloem unloading in sink leaves of barley. Plant Physiol **125:** 209-218

Heldt HW, Chon CJ, Maronde D (1977) Role of orthophosphate and other factors in the regulation of starch formation in leaves and isolated chloroplasts. Plant Physiol **59:** 1146-1155

Hill J, Nelson E, Tilman D, Polasky S, Tiffany D (2006) Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. Proc Natl Acad Sci **103:** 11206-11210

Hirose T, Imaizumi N, Scofield GN, Furbank RT, Ohsugi R (1997) cDNA cloning and tissue specific expression of a gene for sucrose transporter from rice (*Oryza sativa* L.). Plant Cell Physiol **38:** 1389-1396

Hirose T, Takano M, Terao T (2002) Cell wall invertase in developing rice caryopsis: molecular cloning of *OsCIN1* and analysis of its expression in relation to its role in grain filling. Plant Cell Physiol **43:** 452-459

Hofmann K, Stoffel W (1993) A database of membrane spanning protein segments. Biol Chem Hoppe-Seyler **374:** 166

Ishimaru K, Hirose T, Aoki N, Takahashi S, Ono K, Yamamoto S, Wu J, Saji S, Baba T, Ugaki M, Matsumoto T, Ohsugi R (2001) Antisense expression of a rice sucrose transporter *OsSUT1* in rice (*Oryza sativa* L.). Plant Cell Physiol **42:** 1181-1185

Keeling PL, Wood JR, Tyson RH, Bridges IG (1988) Starch biosynthesis in developing wheat grain. Plant Physiol **87**: 311-319

Kong L, Wang F, Feng B, Li S, Si J, Zhang B (2010) The structural and photosynthetic characteristics of the exposed peduncle of wheat (*Triticum aestivum* L.): an important photosynthate source for grain-filling. BMC Plant Biol **10:** 141

Kriedemann P (1966) The photosynthetic activity of the wheat ear. Ann Bot **30:** 349-363

Kuhn C (2003) A comparison of the sucrose transporter systems of different plant species. Plant Biol **5:** 215-232

Kuhn C (2011) Sucrose transporters and plant development. In M Geisler, K Venema, eds, Transporters and pumps in plant signaling, Springer Heidelberg Dordrecht London New York, pp 225-251

Kuhn C, Barker L, Burkle L, Frommer WB (1999) Update on sucrose transport in higher plants. J Exp Bot **50**: 935-953

Kuhn C, Franceschi VR, Schulz A, Lemoine R, Frommer WB (1997) Macromolecular trafficking indicated by localization and turnover of sucrose transporters in enucleate sieve elements. Science **275:** 1298-1300.

Kuhn C, Grof CPL (2010) Sucrose transporters of higher plants. Curr Opin Plant Biol **13:** 288-298

Lalonde S, Boles E, Hellmann H, Barker L, Patrick JW, Frommer WB, Ward JM (1999) The dual function of sugar carriers: transport and sugar sensing**.** Plant Cell **11**: 707-726

Lalonde S, Wipf D, Frommer WB (2004) Transport mechanisms for organic forms of carbon and nitrogen between source and sink. Annu Rev Plant Biol **55:** 341-72

Leggewie G, Kolbe A, Lemoine R, Roessner U, Lytovchenko A, Zuther E, Kehr J, Frommer WB, Reismeir JW, Willmitzer L, Fernie AR (2003) Overexpression of the sucrose transporter SoSUT1 in potato results in alteration in leaf carbon partitioning and in tuber metabolism but has little impact on tuber morphology. Planta **217:** 158-167

Lemoine R (2000) Sucrose transporters in plants: update on function and structure. Biochim Biophys Acta **1465:** 246-262

Li X, Weng JK, Chapple C (2008) Improvement of biomass through lignin modification. Plant J **54:** 569-581

Li Z, Trick HN (2005) Rapid method for high-quality RNA isolation from seed endosperm containing high levels of starch. BioTechniques **38:** 872

Licht FO (2011) Global Renewable Fuels Alliance (GRFA report). http://www.marketwire.com/press-release/global-ethanol-production-to-reach-887-billion-litres-in-2011-1395312.htm (accessed on 29 September 2011)

Lim JD, Cho J, Park Y, Hahn TR, Choi SB, Jeon SJ (2006) Sucrose transport from source to sink seeds in rice. Physiol Plantarum **126:** 572-584

Livak K, Schmittigen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. Methods **25:** 402-408

Lu J, Sheahan C, Fu P (2011) Metabolic engineering of algae for fourth generation biofuels production. Energy Environ Sci **4:** 2451-2466

Lu JMY, Bush DR (1998) His-65 in the proton–sucrose symporter is an essential amino acid whose modification with site-directed mutagenesis increases transport activity. Proc Natl Acad Sci **95:** 9025-9030

Marger MD, Saier MH (1993) A major superfamily of transmembrane facilitators that catalyse uniport, symport and antiport. Trends Biochem Sci **18:** 13-20

Martin C, Smith AM (1995**)** Starch biosynthesis. Plant Cell **7:** 971-985

Matsuoka Y (2011) Evolution of polyploid triticum wheats under cultivation: The role of domestication, natural hybridization and allopolyploid speciation in their diversification. Plant Cell Physiol **52:** 750-764

Naik SN, Goud VV, Rout PK, Dalai AK (2010) Production of first and second generation biofuels: a comprehensive review. Renew Sust Energ Rev **14:** 578-597

Olai M, Romain R, Bergeron N, Klein K (2004) Ethanol industry in Canada. http://www.ascension-publishing.com/BIZ/HD44.pdf (accessed on 4 October, 2011)

Oparka KJ (1990) What is phloem unloading. Plant Physiol **94**: 393-396

Pao SS, Paulsen IT, Saier MH (1998) Major facilitator superfamily. Microbiol Mol Biol Rev **62:** 1-34

Patrick JW, Offler CE (1995) Post-sieve element transport of sucrose in developing seeds. Aust J Plant Physiol **22:** 681-702

Patrick JW, Offler CE (2001) Compartmentation of transport and transfer events in developing seeds. J Exp Bot **52:** 551-564

Rajendrani M (2011) Generations: biofuels are grouped according to commercial maturity. Chemical and Engineering News, http:// pubs.acs .org/cen/coverstory/89/8933cover1.html (accessed on 12 October 2011)

Reinders A, Schulze W, Kuhn C, Barker L, Schulz A, Ward JM, Frommer WB (2002) Protein–protein interactions between sucrose transporters of different affinities colocalized in the same enucleate sieve element. Plant Cell **14:** 1567-1577

Reinders A, Sivitz AB, Strker CG, Gantt JS, Ward JM (2008) Functional analysis of *LjSUT4*, a vacuolar sucrose transporter from *Lotus japonicus*. Plant Mol Biol **68:** 289-299

Rharrabti Y, Elhani S, Martos Nunez V , Garcia del Moral LF (2001) Protein and lysine content, grain yield, and other technological traits in Durum wheat under Mediterranean conditions. J Agric Food Chem **49:** 3802-3807

Riesmeier JW, Hirner B, Frommer WB (1993) Potato sucrose transporter expression in minor veins indicates a role in phloem loading. Plant Cell **5:** 1591-1598

Riesmeier JW, Willmitzer L, Frommer WB (1992) Isolation and characterization of a sucrose carrier cDNA from spinach by functional expression in yeast. EMBO J **11:** 4705-4713

Riesmeier JW, Willmitzer L, Frommer WB (1994) Evidence for an essential role of the sucrose transporter in phloem loading and assimilate partitioning EMBO J **13:** 1-7

Rolland N, Ferro M, Seigneurin-Berny D, Garin J, Douce R, Joyard J (2003) Proteomics of chloroplast envelope membranes. Photosynth Res **78:** 205-230

Rosche E, Blackmore D, Tegeder M, Richardson T, Schroeder H, Higgins TJV, Frommer WB, Offler CE, Patrick JW (2002) Seed-specific over expression of a potato sucrose transporter increases sucrose uptake and growth rates of developing pea cotyledons. Plant J **30:** 165-175

Sanderson K (2011) Lignocellulose: a chewy problem. Nature **474:** S12-S14

Sauer N (2007) Molecular physiology of higher plant sucrose transporters. FEBS Lett **581:** 2309-2317

Saunders J, Izydorczyk M, Levin DB (2011) Limitations and challenges for wheat-based bioethanol production. In IM Aurelio dos Santos Bernardes, eds, Economic effects of biofuel production, InTech, http://www.intechopen.com/articles/show /title/limitations-and-challenges-for-wheat-based-bioethanol-production (accessed on 29 September 2011)

Schnyder H (1993) The role of carbohydrate storage and redistribution in the source-sink relations of wheat and barley during grain filling - a review. New Phytol **123:** 233-245

Schulze W, Weise A, Frommer WB, Ward JM (2000) Function of the cytosolic N-terminus of sucrose transporter AtSUT2 in substrate affinity. FEBS Lett **485:** 189-194

Scofield GN, Hirose T, Gaudron JA, Upadhyaya NM, Ohsugi R, Furbank RT (2002) Antisense suppression of the rice sucrose transporter gene, *OsSUT1*, leads to impaired grain filling and germination but does not affect photosynthesis. Funct Plant Biol **29:** 815-826

Scofield GN, Hirose T, Aoki N, Furbank RT (2007) Involvement of the sucrose transporter, *OsSUT1*, in the long-distance pathway for assimilate transport in rice. J Exp Bot **58:** 3155-3169

Secan - Canada Seed Partner (2009) Ac Andrew - soft white spring wheat variety information. http://www.secan.com/en/data/files/workspace/secansearchvalues/W007/AC%20Andrew%20Feb%202009%20TH.pdf (accessed on 29 September, 2011)

Shewry PR (2009) Wheat. J Exp Bot **60:** 1537-1553

Shiratake K (2007) Genetics of sucrose transporters in plants. Genes Genome Genom **1:** 73-80

Simmons SR (1987) Growth, development and physiology. In E G Heyne eds, Wheat and wheat improvement, Ed 2, American Society of Agronomy, Wisconsin, USA, pp 92-95

Simpson-Holley M, Evans G (2007) Fuelling: the end of the oil age. Outlooks Pest Manag **18:** 109-113

Sivitz AB, Reinders A, Ward JM (2005) Analysis of the transport activity of barley sucrose transporter HvSUT1. Plant Cell Physiol **46:** 1666-1673

Slewinski TL, Meeley R, Braun DM. (2009) *Sucrose transporter 1* functions in phloem loading in maize leaves. J Exp Bot **60:** 881-892

Smith AM (2008) The prospects for increasing starch and sucrose yields for bioethanol production. Plant J **54:** 546-558

Srinivasan J (2008) Climate change, greenhouse gases and aerosols. Resonance **13:** 1146-1155

Stadler R, Brandner J, Schulz A, Gahrtz M, Sauer N (1995) Phloem loading by the PmSUC2 sucrose carrier from Plantago major occurs into companion cells. Plant Cell **7:** 1545-1554

Stadler R, Sauer N (1996) The *Arabidopsis thaliana AtSUC2* gene is specifically expressed in companion cells. Bot Acta **109:** 299-306

Sun Y, Xu HL, Gong WK, Zhai HL, Meng K, Wang YQ, Wei XL, Xiao GF, Zhi Z (2008) Cloning and expression analysis of rice sucrose transporter genes *OsSUT2M* and *OsSUT5Z.* J Integr Plant Biol **50:** 62-75

Strategic Vision Consulting Ltd. (2007) Cereal competitiveness in western Canada: oppurtunities and threats facing the grain-based biofuel and feed industries in western Canada. http://www.acidf.ca/files/focus/ot.pdf (accessed on 29 September 2011]

Taiz L, Zieger E (2010a) Photosynthesis: the light reactions. In Plant physiology, Ed 5, Sinauer Associates, Sunderland,, Massachusetts, USA, pp 163-198

Taiz L, Zieger E (2010b) Photosynthesis: the carbon reactions. In Plant physiology, Ed 5, Sinauer Associates, Sunderland, Massachusetts, USA, pp 199-242

Taiz L, Zieger E (2010c) Translocation in the phloem. In Plant physiology, Ed 5, Sinauer Associates, Sunderland, Massachusetts, USA, pp 272-303

Takeda T, Toyofuku K, Matsukura C, Yamaguchi J (2001) Sugar transporters involved in flowering and grain development of rice. J Plant Physiol **158:** 465-470

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol **28:** 2731-2739

Tomlinson K, Denyer K (2003) Starch synthesis in cereal grains. Adv Bot Res **40:** 1-61

Trethewey RN, Smith AM (2000) Starch metabolism in leaves. In RC Leegood, TD Sharkey, SV Caemmerer eds, Photosynthesis: physiology and metabolism, Vol 9, pp 205-231

Truernit E (2001) The importance of sucrose transporters. Curr Biol **11:** R169-R171

Truernit E, Sauer N (1995) The promoter of the *Arabidopsis thaliana SUC2* sucrose-H+ symporter gene directs expression of beta-glucuronidase to the phloem: evidence for phloem loading and unloading by *SUC2*. Planta **196:** 564-570

US Energy Information Administration (2011) International energy outlook 2011. http://www.eia.gov/forecasts/ieo/pdf/0484(2011).pdf (accessed on 29 September 2011)

Wang ZM, Wei AL, Zheng DM (2001) Photosynthetic characteristics of non-leaf organs of winter wheat cultivars differing in ear type and their relationship with grain mass per ear. Photosynthetica **39:** 239-244

Weise A, Barker L, Kuhn C, Lalonde S, Buschmann H, Frommer WB, Ward JM (2000) A new subfamily of sucrose transporters, SUT4, with low affinity/high capacity localized in enucleate sieve elements of plants. Plant Cell **12:** 1345-1355

Weschke W, Panitz R, Sauer N, Wany Q, Neubohn B, Weber H, Wobus U (2000) Sucrose transport into barley seeds: molecular characterization of two transporters and implications for seed development and starch accumulation. Plant J **21:** 455-467

Williams LE, Lemoine R, Sauer N (2000) Sugar transporters in higher plants - a diversity of roles and complex regulation. Trends Plant Sci **5:** 283-90

Wolosiuk RA, Pontis HG (1974**)** The role of sucrose and sucrose synthetase in carbohydrate plant metabolism. Mol Cell Biochem **4:** 115-12

Wright DP, Scholes JD, Read DJ, Rolfe SA (2005) European and african maize cultivars differ in their physiological and molecular responses to mycorrhizal infection. New Phytol **167**: 881-896

APPENDIX

Appendix 1 Alignment of the nucleotide sequences of SUT2 genes from barley, rice and maize. Sequences of forward and reverse primers used for amplifying the partial fragment of *TaSUT2* are underlined.

```
HvSUT2          ATGCCGCCGCGCCGGCCCAACACCGGCGGGGGCGGGGGCACGTCGTCGTCGGCGGCGCCG
OsSUT2          ---ATGCCGCGGCGGCCTAGCGGCGGCGGCGGCGG-----CGCTGGTCCCG-CGGCGGCC
ZmSUT2          ---ATGCCGCCGCGC----ACGGCTCCGGCGGCGA-----CCTCAACCCCG-CCGCGG--
                     *****  **      *  *  *** ****      *        ** * ***

HvSUT2          GCGCCGCGCAAGGTCCCTCTCCGGTCGCTGCTCCGGGCGGCGTCGGTGGCGTGCGGGGTG
OsSUT2          GCGGTGCGCAAGGTCCCCCTGCGGAAGCTTCTGCGTGCGGCGTCGGTGGCGTGCGGCGTG
ZmSUT2          ---------AAGGTGCCCCTCCGGAAGCTGCTGCGTGCGGCGTCGGTCGCCTGCGGGGTG
                         ***** ** ** ***  *** ** ** *********** ** ***** ***

HvSUT2          CAGTTCGGGTGGGCGCTGCAGCTGTCCCTGCTGACCCCGTACGTCCAGGAGCTGGGCATC
OsSUT2          CAGTTCGGGTGGGCGCCGCAGCTCTCCCTGCTGACCCCCTACGTCCAGGAGCTCGGCATC
ZmSUT2          CAGTTCGGCTGGGCGCTGCAGCTGTCGTTGCTGACCCCGTACGTGCAGGAGCTGGGCATC
                ******** ******* ****** **  ********** ***** ******** ******

HvSUT2          CCGCACGCCTTCGCCAGCCTGGTGTGGCTGTGCGGCCCGCTCTCCGGGCTCCTGGTGCAG
OsSUT2          CCGCACGCCTTCGCCAGCCTCGTCTGGCTCTGCGGCCCCCTCTCCGGCCTCCTCGTCCAG
ZmSUT2          CCGCACGCCTTTGCCAGTCTCGTCTGGCTGTGCGGTCCGCTGTCCGGCCTCCTCGTCCAG
                *********** ***** ** ** ***** ***** ** ** ***** ***** ** ***

HvSUT2          CCCCTGGTGGGCCACCTCTCGGACCGCATCACGCCGGCCAACTCCCCGCTCGGCCGCCGC
OsSUT2          CCCCTCGTCGGCCACCTCTCCGACCGCATCGCCCCCGCCGCCTCCCCCCTCGGCCGCCGC
ZmSUT2          CCCCTCGTCGGCCACCTCTCCGACCGCATCGGCCCCGCCGCTTCGCCGCTCGGGCGCCGC
                ***** ** *********** *********   ** ***   ** ** ***** ******

HvSUT2          CGGCCCTTCATCGCCGCCGGGGCGGCCTCCATCGCCTTCGCCGTGCTCACCGTCGGCTTC
OsSUT2          CGCCCCTTCATCGCCGCCGGCGCCGCTTCCATCGCTGCCGCAGTTCTCACCGTCAGATTC
ZmSUT2          AGGCCCTTCATCGCCGCCGGCGCCGCGTGCATCGCCGCAGCCGTGCTCACCGTCGGCTTC
                 * ***************** ** ** * ******    ** ** ********* * ***

HvSUT2          TCCGCCGACCTCGGCCGCCTCTTCGGGGACAACGTGGTGCCTGGCTCCACACGCATCGGC
OsSUT2          TCCGCCGACCTCGGCCGAATCTTCGGCGATTCCATCACCCCGGGCTCCACCCGCCTCGGC
ZmSUT2          TCCGCTGACCTCGGCCGACTCTTCGGCGACGACGTCACCCCGGGCTCAACGCGCCTCGGC
                ***** ***********  ******* **   * *    ** ***** ** *** *****

HvSUT2          GCCATCATCGTCTACCTCGTCGGCTTCTGGCTGCTCGACGTCGGCAACAACGCCACCCAG
OsSUT2          GCCATCACCGCCTACCTCGTCGGCTTCTGGCTCCTCGACGTCGGCAACAACGCTACACAG
ZmSUT2          GCCATCTGCGTCTACCTTGTAGGATTCTGGCTGCTCGACGTGGGCAACAACGGCACGCAG
                ******  ** ****** ** ** ******** ******** **********  ** ***

HvSUT2          GGCCCATGCCGCGCCTTCCTCGCCGACCTCACCGAGAATGACCCGAGGAGGACCCGGATC
OsSUT2          GGACCCTGCAGGGCCTTCCCCGCCGACCTCACCGAGAATGACCCAAAGAGGACTCGGATA
ZmSUT2          GGGCCCTGCAGGGCGTTCCTCGCCGACCTCACAGAGAATGACCCAAGGAGGACTCGGATC
                ** ** *** * ** **** ************ *********** * ****** *****

HvSUT2          GCCAATGCCTACTTCTCACTCTTCATGGCCCTGGGGAACATACTCGGGTACGCCACCGGG
OsSUT2          GCTAATGCTTACTTCTCATTGTTCATGGCCCTGGGAAACATACTTGGATATGCCACTGGG
ZmSUT2          GCTAATGCATACTTTTCACTCTTCATGGCCCTGGGAAACATACTTGGATATGCCACCGGA
                ** ***** ***** *** * ************** ******** ** ** ***** **

HvSUT2          GCGTACAATGGCTGGTACAAGATATTCCCGTTCACTATCACTGGGTCCTGTGGCGTCAGC
OsSUT2          GCATACAGTGGCTGGTACAAGATATTCCCGTTCACCGTTACTCCATCATGTAGCATCAGC
ZmSUT2          GCATACAGTGGATGGTATTCGATATTTCCTTTCACTGTTACAGAGTCCTGCGGCATCAGT
                ** **** *** *****   ****** ** *****  * **    ** **  ** ****
```

Continued

```
HvSUT2          TGCGCCAACCTCAATTCTGCGTTCCTGCTTGATATCATCATCCTGGCGATCACGACGTAC
OsSUT2          TGTGCCAACCTCAAGTCGGCCTTTCTACTTGATATTATCATTTTGGTGGTCACTACATGC
ZmSUT2          TGTGCCAACCTTAAGTCTGCCTTTCTTCTTGACATTATTGTTCTGGTCATCACAACGTAC
                ** ******** ** ** ** ** ** ***** ** **  *  ***   **** ** * *

HvSUT2          ATTAGCGTGGCGACGGTGCAGGACAATCCTACATTTGGAAGTGACGAGGCAGCGCCTCCA
OsSUT2          ATCACTGTAGCATCAGTGCAAGAGCCTCAATCCTTTGGAAGTGATGAAGCAGATCACCCT
ZmSUT2          ACTACTGTAACATCCGTGCAAGAACCTCAAACTTTTGGAAGTGATGAAGCACAAAATCCA
                *  *  **  *  * ***** **   **   * ********** ** ***      **

HvSUT2          AGCAGCCACGAAGAGGAAGCTTTCCTCTTTGAGCTATTCGGGTCGTTCAAATACTTCACG
OsSUT2          AGCA---CAGAACAGGAAGCTTTCCTCTGGGAACTTTTTGGATCATTCCGGTACTTTACA
ZmSUT2          GGTG---CTGAACAGGAAGCTTTCCTCTGGGAACTTTTTGGGTCATTAAGATACTTCACC
                 *       *** ***************  ** ** ** ** ** **    ***** **

HvSUT2          ATGCCTGTTTGGATGGTCTTGATCGTCACTTCTCTTACTTGGGTCGGGTGGTTCCTTTTT
OsSUT2          TTACCGGTTTGGATGGTTTTGATTGTTACTGCCCTCACATGGATTGGATGGTTTCCATTT
ZmSUT2          CTACCAATTTGGATGGTCTTAATTGTCACTGCCCTTACTTGGATGGCATGGTTTCCTTTT
                 * **  ********** ** ** ** *** * ** ** *** * *  ***** *  ***

HvSUT2          ATCCTCTTTGACACCGACTGGATGGGCCGAGAGATCTACCGGGGAAGCCCGGAGATCGTC
OsSUT2          ATCCTCTTTGATACCGATTGGATGGGTCGAGAGATCTATCGTGGAAGTCCAGATGATCCA
ZmSUT2          ACCCTCTTTGATACTGATTGGATGGGCCGAGAAATCTATCGAGGAAGCCCAGACAACCCA
                * ********* ** ** ******** ***** ***** ** ***** ** **

HvSUT2          GCCGACACCCAAAAGTATCATGACGGTGTGAGAATGGGCTCTTTTGGTCTCATGCTCAAC
OsSUT2          AGTATAACTCAGAGCTATCATGATGGTGTGAGAATGGGTTCTTTTGGTCTGATGCTGAAC
ZmSUT2          GGAGAGACCCAAAGGTACCATGATGGTGTGAGAATGGGTTCTTTTGGTCTAATGCTCAAC
                      ** ** *  ** ***** ************** *********** ***** ***

HvSUT2          TCGGTCGTTCTCGGGATCACATCTATTGGAATGGAGAAGTTGTGTAGGAAGTGGGGAGCT
OsSUT2          TCGGTCCTTCTTGGATTCACTTCTATTGTACTAGAGAAGTTATGTCGGAAGTGGGGAGCT
ZmSUT2          TCTGTCGTCCTTGGATTCACCTCTGTTGTGCTAGAAAAGTTATGTAGGAAGTGGGGAGCT
                ** *** * ** **  **** *** ***   * ** ***** *** **************

HvSUT2          GGACTTGTATGGGGTGTCTCCAATATCATCATGGCTCTGTGCTTCGTGGCGATGCTCATT
OsSUT2          GGACTGGTGTGGGGTGTCTCCAATATCCTAATGGCATTGTGCTTTGTGGCAATGCTTGTA
ZmSUT2          GGACTTGTATGGGGTGTCTCCAATATTCTAATGACTTTGTGCTTCTTGGCAATGCTTGTA
                ***** ** *****************  * *** *  *******  **** *****  *

HvSUT2          ATAACATACGTGGCGCAGAACCTGGATTATGGACCCAGCGGAGCACCTCCAACCGGAATT
OsSUT2          ATAACATATGTGGCAAAGAATATGGATTATCCACCTAGTGGAGTACCACCAACCGGCATT
ZmSUT2          ATAACATATGTTGCAAAGAACATGGATTATCCATCCAGTGGTGCACCTCCGACTGGCATT
                ******** ** **  ****  ********  * * ** ** * *** ** ** ** ***

HvSUT2          GTCGCCGCTTCCCTCATTGTTTTCACAATCTTAGGAGCACCTCTGTCGATCACGTACAGT
OsSUT2          GTCATTGCTTCCCTGGTAGTTTTTACAATTTTAGGAGCGCCCCTGGCGATCACGTACAGT
ZmSUT2          GTCGTTGCTTCTTTGGTGGTTTTTACAATTTTAGGGGCACCCCTGGCGATCACATACAGT
                ***   *****  *  * ***** ***** ***** ** ** *** ******* ******

HvSUT2          ATACCGTATGCGATGGCTGCAAGTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCA
OsSUT2          ATACCATATGCAATGGCTGCTAGTCGGGTTGAAAATCTGGGACTTGGCCAAGGTCTAGCA
ZmSUT2          ATACCATATGCAATGGCTGCTAGTCGTGTTGAAAATCTCGGCCCTGGCCAAGGCCTAGCA
                ***** ***** ******** ***** *********** ** *  ******** ******

HvSUT2          ATGGGCATTCTTAATTTATCTATTGTCATACCACAGATCATCGTGTCGCTGGGCAGCGGG
OsSUT2          ATGGGCATTCTTAATTTGGCTATTGTCATACCACAGGTTATTGTGTCACTGGGTAGCGGG
ZmSUT2          ATGGGTATTCTCAATTTAGCTATTGTCATACCACAGGTTATTGTGTCACTTGGCAGTGGC
                ***** ***** *****  ***************** * ** ***** ***** ** **
```

Continued

```
HvSUT2          CCGTGGGACCAGCTCTTCGGCGGAGGGAACGCGCCGTCCTTCTGGGTGGCTGCCGCGGCC
OsSUT2          CCCTGGGACCAACTGTTTGGTGGTGGCAATGCACCAGCCTTTGCAGTGGCTGCTGCTGCA
ZmSUT2          CCATGGGATCAGCTGTTTGGCGGTGGAAATGCACCCGCTTTTGCGGTGGCAGCCGGTGCA
                ** ***** ** ** ** ** ** ** ** ** ** * **    ***** ** *  **

HvSUT2          TCATTCGTGGGCGGGCTGGTAGCCATCCTGGGGCTCCCGCGGGCCCGGCTCGGGCCG---
OsSUT2          TCTTTTATCGGTGGGCTGGTGGCTATTCTGGGCCTTCCACGAGCCCGCATTGCATCA---
ZmSUT2          TCTTTCATCGGTGGGCTGGTGGCTATTCTGGGGCTTCCTCGGGCACGCATTGCATCCTCG
                ** **  * ** ******** ** ** ***** ** ** ** ** **  * *   *

HvSUT2          ------AAGAAGAAAACCACCCAACGATGA---
OsSUT2          ------AGGAGGAGAGGT------CACCGATAA
ZmSUT2          TCCTCCAGGAGGAGAGGCGGGACACACCGATAA
                     * ** ** *         *  **
```

Appendix 2 Nucleotide sequence similarity between the partial DNA fragment of *TaSUT2* (query) and other *SUT2* from the GenBank (subject).

EU693244.1 *Triticum aestivum* cv. Bobwhite *sucrose transporter 2-like* (*SUT2*) mRNA, complete sequence, length = 1843

Score = 1421 bits (769), Expect = 0.0
Identities = 796/809 (98%), Gaps = 1/809 (0%)
Strand = Plus/Plus

```
Query  1     CATGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGAT  60
             ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Sbjct  706   CATGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCGCACAACGGCTGGTACAAGAC 765

Query  61    ATTCCCCTTCACTATTACGGAGTCCTGCTCTGTCAGCTGCGCCAACCTCAAGTCCGCGTT 120
             ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct  766   ATTCCCCTTCACTATTACGGAGTCCTGCTCTGTCAGCTGCGCCAACCCCAAGTCCGCGTT 825

Query  121   CCTGCTCGACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGTGACGGTGCAGGA 180
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct  826   CCTGCTCGACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGGGACGGTGCAGGA 885

Query  181   CAATCCTACTTTCGGAGGCGACGAGGCGGCACCTCCGAGCAGCCACGAGGAGGAGGCTTT 240
             ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  886   CAATCCTACTTTCGGAAGCGACGAGGCGGCACCTCCGAGCAGCCACGAGGAGGAGGCTTT 945

Query  241   CCTCTTTGAGCTTTTCGGGTCGTTCAAGTACTTCACGATGCCTGTTTGGATGGTCTTGAT 300
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Sbjct  946   CCTCTTTGAGCTTTTCGGGTCGTTCAAGCACTTCACGATGCCTGTTTGGATGGTCTTGAT1005

Query  301   CGTCACCTCGCTTACCTGGATCGGCTGG-TTCCCTTTCATCCTCTTTGACACCGACTGGA 359
             |||||||||||||||||||||||| ||| ||||||||||||||||||||||||||||||||
Sbjct  1006  CGTCACCTCGCTTACCTGGATCGGGTGGTTTCCCTTTCATCCTCTTTGACACCGACTGGA1065

Query  360   TGGGCCGAGAGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAAAAGTATCATG 419
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1066  TGGGCCGAGAGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAAAAGTATCATG1125

Query  420   ACGGTGTGAGGATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCTCGGGATCACAT 479
             |||||||||| |||||| ||||||||||||||||||||||||||||||||||||||||||
Sbjct  1126  ACGGTGTGAGAATGGGCCCTTTTGGTCTCATGCTCAACTCGGTTGTTCTCGGGATCACAT1185

Query  480   CTATTGGAATGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTATGGGGTGTCTCCA 539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1186  CTATTGGAATGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTATGGGGTGTCTCCA1245

Query  540   ATATCATCATGGCACTGTGCTTCGTGGCGATGCTTATTATAACAAACGTTGCGCAGAACC 599
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Sbjct  1246  ATATCATCATGGCACTGTGCTTCGTGGCGATGCTTATTATAACATACGTTGCGCAGAACC1305

Query  600   TGGATTATGGACCTAGTGGAGCACCTCCAACCGGCATTGTCGCCGCTTCCCTCATTGTGT 659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1306  TGGATTATGGACCTAGTGGAGCACCTCCAACCGGCATTGTCGCCGCTTCCCTCATTGTGT1365

Query  660   TCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTATACCATATGCGATGGCTGCAA 719
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Sbjct  1366  TCACAATCCTAGGAGCACCCCTGTCGGTTACGTACAGTATACCATATGCGATGGCTGCAA1425

Query  720   GTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGCATTCTTAATTTATCTA 779
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1426  GTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGCATTCTTAATTTATCTA1485

Query  780   TTGTCATACCACAGATCATCGTGTCGCTG  808
             |||||||||||||||||||||||||||||
Sbjct  1486  TTGTCATACCACAGATCATCGTGTCGCTG  1514
```

AJ272308.1 *Hordeum vulgare sucrose transporter 2* (*SUT2*) mRNA, complete sequence, length = 2018

Score = 1216 bits (658), Expect = 0.0
Identities = 762/812 (94%), Gaps = 8/812 (1%)
Strand = Plus/Plus

```
Query  1     CATGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGAT  60
             |||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||||||
Sbjct  694   CATGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCGTACAATGGCTGGTACAAGAT  753

Query  61    ATTCCCCTTCACTATTAC-GGAGTCCTGCT--CTGTCAGCTGCGCCAACCTCAAGTCCGC  117
             |||||| |||||||| || || |||||| |  | ||||||||||||||||||| || ||
Sbjct  754   ATTCCCGTTCACTATCACTGG-GTCCTG-TGGC-GTCAGCTGCGCCAACCTCAATTCTGC  810

Query  118   GTTCCTGCTCGACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGTGACGGTGCA  177
             ||||||||| || ||||||||||||||||| ||||||||| ||||||| | |||||||||
Sbjct  811   GTTCCTGCTTGATATCATCATCCTGGCGATCACGACGTACATTAGCGTGGCGACGGTGCA  870

Query  178   GGACAATCCTAC-TTTCGGAGGCGACGAGGCGGCACCTCCGAGCAGCCACGAGGAGGAGG  236
             |||||||||||| ||| ||| | |||||||| || ||||| |||||||||||| ||||| |
Sbjct  871   GGACAATCCTACATTT-GGAAGTGACGAGGCAGCGCCTCCAAGCAGCCACGAAGAGGAAG  929

Query  237   CTTTCCTCTTTGAGCTTTTCGGGTCGTTCAAGTACTTCACGATGCCTGTTTGGATGGTCT  296
             ||||||||||||||| ||||||||||||| |||||||||||||||||||||||||||||
Sbjct  930   CTTTCCTCTTTGAGCTATTCGGGTCGTTCAAATACTTCACGATGCCTGTTTGGATGGTCT  989

Query  297   TGATCGTCACCTCGCTTACCTGGATCGGCTGGTTCCCTTTCATCCTCTTTGACACCGACT  356
             |||||||||| || ||||| ||| |||| ||||||| ||| ||||||||||||||||||
Sbjct  990   TGATCGTCACTTCTCTTACTTGGGTCGGGTGGTTCCTTTTTATCCTCTTTGACACCGACT  1049

Query  357   GGATGGGCCGAGAGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAAAAGTATC  416
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1050  GGATGGGCCGAGAGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAAAAGTATC  1109

Query  417   ATGACGGTGTGAGGATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCTCGGGATCA  476
             |||||||||||||| |||||||||||||||||||||||||||| |||||||||||||||
Sbjct  1110  ATGACGGTGTGAGAATGGGCTCTTTTGGTCTCATGCTCAACTCGGTCGTTCTCGGGATCA  1169

Query  477   CATCTATTGGAATGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTATGGGGTGTCT  536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1170  CATCTATTGGAATGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTATGGGGTGTCT  1229

Query  537   CCAATATCATCATGGCACTGTGCTTCGTGGCGATGCTTATTATAACAAACGTTGCGCAGA  596
             |||||||||||||||| |||||||||||||||||| ||||||||| |||| |||||||
Sbjct  1230  CCAATATCATCATGGCTCTGTGCTTCGTGGCGATGCTCATTATAACATACGTGGCGCAGA  1289

Query  597   ACCTGGATTATGGACCTAGTGGAGCACCTCCAACCGGCATTGTCGCCGCTTCCCTCATTG  656
             |||||||||||||||| || |||||||||||||||||| ||||||||||||||||||||||
Sbjct  1290  ACCTGGATTATGGACCCAGCGGAGCACCTCCAACCGGAATTGTCGCCGCTTCCCTCATTG  1349

Query  657   TGTTCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTATACCATATGCGATGGCTG  716
             | |||||||| |||||||||||||||| | |||||||||||||| ||||||||||||||
Sbjct  1350  TTTTCACAATCTTAGGAGCACCTCTGTCGATCACGTACAGTATACCGTATGCGATGGCTG  1409

Query  717   CAAGTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGCATTCTTAATTTAT  776
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1410  CAAGTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGCATTCTTAATTTAT  1469

Query  777   CTATTGTCATACCACAGATCATCGTGTCGCTG  808
             ||||||||||||||||||||||||||||||||
Sbjct  1470  CTATTGTCATACCACAGATCATCGTGTCGCTG  1501
```

GU045300.1 *Sorghum bicolor sucrose transporter 4* (*SUT4*) mRNA, complete sequence, length = 1835

Score = 518 bits (280), Expect = 3e-143
Identities = 641/815 (79%), Gaps = 25/815 (3%)
Strand = Plus/Plus

```
Query  1     CATGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGAT  60
             ||||||||||||  ||||||||  ||  ||  ||||||||  |||||||   ||  |||||   |||
Sbjct  576   CATGGCCCTGGGAAACATACTTGGATATGCCACCGGAGCATACAGTGGATGGTATTCGAT  635

Query  61    ATTCCCCTTCACTATTACGGAGTCCTGCTCTGTCAGCTGCGCCAACCTCAAGTCCGCGTT  120
             |||  ||  ||||||  ||||  |||||||||    |||||  ||  ||||||||  |||||  ||  ||
Sbjct  636   ATTTCCTTTCACTGTTACAGAGTCCTGCGGAGTCAGTTGTGCCAACCTTAAGTCTGCCTT  695

Query  121   CCTGCTCGACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGT-G-A-CGGTGCA  177
              ||  ||  |||||  ||  ||  ||||  ||||||  ||||||  |  |  |  |  ||  |  |  |  |||||
Sbjct  696   TCTTCTTGACATTATTATTCTGGTGATTACAACGTACATCA-C-T-GTAGCATCCGTGCA  752

Query  178   GGA-CAATCCTACTTTCGGAGGCGACGAGGCGGCACCTCCGAGCAGCCACGAGGAGGAGG  236
              ||  |  |  |  |||||  |||    ||  ||  ||   ||    ||  ||   |  |   ||   ||||  |
Sbjct  753   AGAGCCA-CAAACTTTTGGAAATGATGAAGC-ACAAAACCCAGGTG-C-TGAACAGGAAG  808

Query  237   CTTTCCTCTTTGAGCTTTTCGGGTCGTTCAAG-TACTTCACGATGCCTGTTTGGATGGTC  295
             |||||||||   ||  |||||  |||||  ||  |||  ||||||||   |  ||   ||||||||||||
Sbjct  809   CTTTCCTCTGGGAACTTTTTGGGTCATT-AAGATACTTCACCCTACCAATTTGGATGGTC  867

Query  296   TTGATCGTCACCT-CGCTTACCTGGATCGGC-TGGTTCCCTTTCATCCTCTTTGACACCG  353
             ||  ||  ||||  ||  |  |||||  ||||||  ||  |||||  |||||  |  |||||||||  ||  |
Sbjct  868   TTAATTGTCA-CTGCCCTTACGTGGATC-GCATGGTTTCCTTTTACCCTCTTTGATACTG  925

Query  354   ACTGGATGGGCCGAGAGATCTACCGGGGAAGCCCGGAGATCGTC-GCCGACACCCAAAAG  412
             |  |||||||||||||  ||||||||  ||||||||  ||  |  |   |  |   ||  |||||||  |
Sbjct  926   ATTGGATGGGCCGAGAAATCTACCGAGGAAGCCCAGACAAC-CCAGGAGAGACCCAAAGG  984

Query  413   TATCATGACGGTGTGAGGATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCTCGGG  472
             ||||||||  ||||||||||||||  |||||||||||||||||||||||  ||  ||  ||  ||
Sbjct  985   TATCATGATGGTGTGAGGATGGGTTCTTTTGGTCTCATGCTCAACTCTGTCGTCCT-TGG  1043

Query  473   A-TCACATCTATTG-GAATGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTATGGG  530
             |  ||||  |||  |||  |   |  ||  |||||  ||||||||||||||||||||||||||||||||
Sbjct  1044  ATTCACCTCTGTTGTG-CTAGAAAAGTTATGTAGGAAGTGGGGAGCTGGACTTGTATGGG  1102

Query  531   GTGTCTCCAATATCATCATGGCACTGTGCTTCGTGGCGATGCTTATTATAACAAACGTTG  590
             ||||||||||||  ||  |||  |  ||||||||  ||||  ||||||  |  ||||||  |  ||||
Sbjct  1103  GTGTCTCCAATATTATAATGACTTTGTGCTTCTTGGCAATGCTTGTAATAACATATGTTG  1162

Query  591   CGCAGAACCTGGATTATGGA-CCTAGTGGAGCACCTCCAACCGGCATTGTCGCCGCTTCC  649
             |  ||||||  ||||||||  |  ||  |  |||  |  ||||||  ||  ||||||  |||  ||  ||
Sbjct  1163  CACAGAACATGGATTATCCATCC-ACTGGTGAACCTCCGACTGGCATTATCGTTGCCTCT  1221

Query  650   CTCATTGTGTTCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTATACCATATGCG  709
              |  |  ||  ||  ||  ||  ||||  |||||  |||  |  |  ||  ||||||||||||||||||
Sbjct  1222  TTGGTGGTTTTTACGATTTTAGGGGCACCCCTGGCAATCACATACAGTATACCATATGCA  1281

Query  710   ATGGCTGCAAGTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGCATTCTT  769
             ||||||||  ||||||||||||||||||  ||  ||||||||||||||||||||  |||||
Sbjct  1282  ATGGCTGCTAGTCGTGTTGAAAATCTTGGTCTTGGCCAAGGTCTAGCAATGGGTATTCTC  1341

Query  770   AATTTATCTATTGTCATACCACAGATCATCGTGTC  804
             ||||||  ||||||||  ||||||||  |  ||  |||||
Sbjct  1342  AATTTAGCTATTGTCGTACCACAGGTTATTGTGTC  1376
```

AB091672.1 *Oryza sativa* Japonica Group *sucrose transporter 2* (*SUT2*) mRNA, complete sequence, length = 1869

Score = 508 bits (275), Expect = 2e-140
Identities = 648/825 (79%), Gaps = 37/825 (4%)
Strand = Plus/Plus

```
Query  1     CATGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGAT  60
             ||||||||||||  ||||||||  ||  ||  |||||  ||||||||||   ||||||||||||||
Sbjct  630   CATGGCCCTGGGAAACATACTTGGATATGCCACTGGGGCATACAGTGGCTGGTACAAGAT  689

Query  61    ATTCCCCTTCACTATTAC-GGAGTCCTGCTCTG--TCAGCTGCGCCAACCTCAAGTCCGC  117
             ||||||  |||||   ||||    | ||  ||  |   |   |||||||  ||||||||||||||  ||
Sbjct  690   ATTCCCGTTCACCGTTACTCCA-TCATG-T-AGCATCAGCTGTGCCAACCTCAAGTCGGC  746

Query  118   GTTCCTGCTCGACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGT-G-A-CGGT  174
              ||  ||  ||  ||  ||  |||||   |||  |  |  ||  ||  |  |  |  |  |  ||  |  |  |  ||
Sbjct  747   CTTTCTACTTGATATTATCATTTTGGTGGTCACTACATGCATCA-C-T-GTAGCATCAGT  803

Query  175   GCAGGA-CAATC-CTACTTTCGGAGGCGACGAGGC-G-GCACCTCCGAGCAGCCACGAGG  230
             |||  ||  |   ||   |  ||||  |||  |  ||  ||  ||  |   ||  |  ||  ||||   ||  ||
Sbjct  804   GCAAGAGC-CTCAATCCTTT-GGAAGTGATGAAGCAGATCA-C-CCTAGCA--CA-GAAC  856

Query  231   AGGAGGCTTTCCTCTTTGAGCTTTTCGGGTCGTTCAAGTACTTCACGA-TGCCTGTTTGG  289
             ||||  ||||||||||   ||  |||||  ||  ||  |||   ||||||  ||  |  |  ||  ||||||
Sbjct  857   AGGAAGCTTTCCTCTGGGAACTTTTTGGATCATTCCGGTACTTTAC-ATTACCGGTTTGG  915

Query  290   ATGGTCTTGATCGTCACCT-CGCTTACCTGGATCGGCTGG-TTCCCTTTCATCCTCTTTG  347
             |||||  |||||  ||  |  ||  |  ||  ||  |||||  ||  |||  ||||  |||  ||||||||||
Sbjct  916   ATGGTTTTGATTGTTA-CTGCCCTCACATGGATTGGATGGTTTCCATTT-ATCCTCTTTG  973

Query  348   ACACCGACTGGATGGGCCGAGAGATCTACCGGGGAAGCCCGGA-GATCGTCGCCGACACC  406
             |  |||||  ||||||||  ||||||||||  ||  |||||  ||  ||  ||||   |   |  ||
Sbjct  974   ATACCGATTGGATGGGTCGAGAGATCTATCGTGGAAGTCCAGATGATCCAAGTATA-ACT  1032

Query  407   CAAAAG-TATCATGACGGTGTGAGGATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGT  465
             |  |  ||  ||||||||  ||||||||  |||||  ||||||||||  |||||  ||||||||   |
Sbjct  1033  C-AGAGCTATCATGATGGTGTGAGAATGGGTTCTTTTGGTCTGATGCTGAACTCGGTCCT  1091

Query  466   TCTCGGGA-TCACATCTATTGGAATGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTG  524
             |||  |||  ||||  |||||||  |  |  ||||||||  |||  ||||||||||||||||||  |
Sbjct  1092  TCT-TGGATTCACTTCTATTGTACTAGAGAAGTTATGTCGGAAGTGGGGAGCTGGACTGG  1150

Query  525   TATGGGGTGTCTCCAATATCATCATGGCACTGTGCTTCGTGGCGATGCTTATTATAACAA  584
             |  ||||||||||||||||||  |  ||||||  |||||||  |||||  ||||||  |  ||||||
Sbjct  1151  TGTGGGGTGTCTCCAATATCCTAATGGCATTGTGCTTTGTGGCAATGCTTGTAATAACAT  1210

Query  585   ACGTTGCGCAGAACCTGGATTATGGACCTAGTGGAGCACCTCCAACCGGCATTGTCGCCG  644
             |  ||  ||   ||||   ||||||||   ||||||||||  |||  ||||||||||||   |
Sbjct  1211  ATGTGGCAAAGAATATGGATTATCCACCTAGTGGAGTACCACCAACCGGCATTGTCATTG  1270

Query  645   CTTCCCTCATTGTGTTCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTATACCAT  704
             |||||||  |  ||  ||  |||||  |||||||  ||  |||  ||  |  ||||||||||||||||
Sbjct  1271  CTTCCCTGGTAGTTTTTACAATTTTAGGAGCGCCCCTGGCGATCACGTACAGTATACCAT  1330

Query  705   ATGCGATGGCTGCAAGTCGTGTTGAAAATCTTGGG-CTAGGCCAAGGTCTAGCAATGGGC  763
             ||||  ||||||||  |||||  ||||||||||  ||||  ||  ||||||||||||||||||||
Sbjct  1331  ATGCAATGGCTGCTAGTCGGGTTGAAAATC-TGGGACTTGGCCAAGGTCTAGCAATGGGC  1389

Query  764   ATTCTTAATTTATCTATTGTCATACCACAGATCATCGTGTCGCTG  808
             ||||||||||  ||||||||||||||||  |  ||  |||||  |||
Sbjct  1390  ATTCTTAATTTGGCTATTGTCATACCACAGGTTATTGTGTCACTG  1434
```

AY639018.1 *Zea mays sucrose transport 4* (*SUT4*) mRNA, complete sequence, length = 1649

Score = 501 bits (271), Expect = 3e-138
Identities = 643/820 (78%), Gaps = 35/820 (4%)
Strand = Plus/Plus

```
Query  1     CATGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGAT  60
             ||||||||||||  ||||||||  ||  ||  ||||||||  |||||||   ||  |||||   |||
Sbjct  659   CATGGCCCTGGGAAACATACTTGGATATGCCACCGGAGCATACAGTGGATGGTATTCGAT  718

Query  61    ATTCCCCTTCACTATTACGGAGTCCTGC-TCTGTCAGCTGCGCCAACCTCAAGTCCGCGT  119
             |||  ||  ||||||  ||||  |||||||||   |   ||||  ||  ||||||||  |||||  ||  |
Sbjct  719   ATTTCCTTTCACTGTTACAGAGTCCTGCGGC-ATCAGTTGTGCCAACCTTAAGTCTGCCT  777

Query  120   TCCTGCTCGACATCATCATCCTGG-CGATTACGACGTACGTTAGCGTCGT--GA-CGGTG  175
             |  ||  ||  |||||  ||   |  ||||  |  ||  ||  ||||||   ||  |  |  ||    |  |  |||
Sbjct  778   TTCTTCTTGACATTATTGTTCTGGTC-ATCACAACGTACACTA-C-T-GTAACATCCGTG  833

Query  176   C-AGGACAATCCTACTTTCGGAAGCGACGAGGCGGC-ACCTCCGAGCAGCCACGAGGAGG  233
             |  ||  ||   ||   |||||  |||||  ||  ||  ||   |  |   |||  ||   |  |   ||   |||
Sbjct  834   CAAGAAC-CTCAAACTTTTGGAAGTGATGAAGC-ACAAAATCC-AGGTG-C-TGAACAGG  888

Query  234   AGGCTTTCCTCTTTGAGCTTTTCGGGTCGTTCAAG-TACTTCACGATGCCTGTTTGGATG  292
             |  ||||||||||   ||  |||||  |||||  ||  |||  ||||||||   |  ||   ||||||||
Sbjct  889   AAGCTTTCCTCTGGGAACTTTTTGGGTCATT-AAGATACTTCACCCTACCAATTTGGATG  947

Query  293   GTCTTGATCGTCACCT-CGCTTACCTGGATCGGC-TGGTTCCCTTTCATCCTCTTTGACA  350
             |||||  ||  ||||  ||  |  |||||  |||||  |||  |||||  |||||  |  |||||||||  |
Sbjct  948   GTCTTAATTGTCA-CTGCCCTTACTTGGAT-GGCATGGTTTCCTTTTACCCTCTTTGATA  1005

Query  351   CCGACTGGATGGGCCGAGAGATCTACCGGGGAAGCCCGGAGATCGTC-GCCGACACCCAA  409
             |  ||  ||||||||||||||  |||||  ||  ||||||||  ||  |  |   |  |   ||  ||||||
Sbjct  1006  CTGATTGGATGGGCCGAGAAATCTATCGAGGAAGCCCAGACAAC-CCAGGAGAGACCCAA  1064

Query  410   AAGTATCATGACGGTGTGAGAATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCTC  469
             |  |||  |||||  |||||||||||||||  ||||||||||  ||||||||||  ||  ||  ||
Sbjct  1065  AGGTACCATGATGGTGTGAGAATGGGTTCTTTTGGTCTAATGCTCAACTCTGTCGTCCT-  1123

Query  470   GGGA-TCACATCTATTG-GAATGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTAT  527
              |||  ||||  |||  |||  |   |  ||  |||||  ||||||||||||||||||||||||||||||
Sbjct  1124  TGGATTCACCTCTGTTGTG-CTAGAAAAGTTATGTAGGAAGTGGGGAGCTGGACTTGTAT  1182

Query  528   GGGGTGTCTCCAATAT-CATCATGGCACTGTGCTTCGTGGCGATGCTTATTATAACATAC  586
             ||||||||||||||||  |  |  |||  |   ||||||||  ||||  ||||||  |  ||||||||
Sbjct  1183  GGGGTGTCTCCAATATTC-TAATGACTTTGTGCTTCTTGGCAATGCTTGTAATAACATAT  1241

Query  587   GTTGCGCAGAACCTGGATTATGGA-CCTAGTGGAGCACCTCCAACCGGCATTGTCGCCGC  645
             |||||  |||||  ||||||||  |  ||  |||||  ||||||||  ||  ||||||||||  ||
Sbjct  1242  GTTGCAAAGAACATGGATTATCCATCC-AGTGGTGCACCTCCGACTGGCATTGTCGTTGC  1300

Query  646   TTCCCTCATTGTGTTCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTATACCATA  705
             |||  |  |  ||  ||  |||||  ||||  |||||  |||  ||  |  ||  ||||||||||||||
Sbjct  1301  TTCTTTGGTGGTTTTTACAATTTTAGGGGCACCCCTGGCGATCACATACAGTATACCATA  1360

Query  706   TGCGATGGCTGCAAGTCGTGTTGAAAATCTTGGGC-TAGGCCAAGGTCTAGCAATGGGCA  764
             |||  ||||||||  ||||||||||||||||||  ||  |  |  ||||||||  ||||||||||  |
Sbjct  1361  TGCAATGGCTGCTAGTCGTGTTGAAAATCTCGGCCCT-GGCCAAGGCCTAGCAATGGGTA  1419

Query  765   TTCTTAATTTATCTATTGTCATACCACAGATCATCGTGTC  804
             ||||  ||||||  ||||||||||||||||||  |  ||  |||||
Sbjct  1420  TTCTCAATTTAGCTATTGTCATACCACAGGTTATTGTGTC  1459
```

Appendix 3 Nucleotide sequence of the partial fragment of *TaSUT2*. The nested primer sets used for 5' RACE are underlined.

CATGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGA
TATTCCCCTTCACTATTACGGAGTCCTGCTCTGTCAGCTGCGCCAACCTCAAGTCCGCG
TTCCTGCTCGACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGTGACGGTGCA
GGACAATCCTACTTTCGGAAGCGACGAGGCGGCACCTCCGAGCAGCCACGAGGAGGAGG
CTTTCCTCTTTGAGCTTTTCGGGTCGTTCAAGTACTTCACGATGCCTGTTTGGATGGTC
TTGATCGTCACCTCGCTTACCTGGATCGGCTGGTTCCCTTTCATCCTCTTTGACACCGA
CTGGATGGGCCGAGAGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAAAAGT
ATCATGACGGTGTGAGAATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCTCGGG
ATCACATCTATTGGAATGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTATGGGG
TGTCTCCAATATCATCATGGCACTGTGCTTCGTGGCGATGCTTATTATAACATACGTTG
CGCAGAACCTGGATTATGGACCTAGTGGAGCACCTCCAACCGGCATTGTCGCCGCTTCC
CTCATTGTGTTCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTATACCATATGC
GATGGCTGCAAGTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGCATTC
TTAATTTATCTATTGTCATACCACAGATCATCGTGTCGCTG

Appendix 4 Nucleotide sequence of the partial fragment of *TaSUT2*. The nested primer sets used for 3' RACE are underlined.

CATGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGA
TATTCCCCTTCACTATTACGGAGTCCTGCTCTGTCAGCTGCGCCAACCTCAAGTCCGCG
TTCCTGCTCGACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGTGACGGTGCA
GGACAATCCTACTTTCGGAAGCGACGAGGCGGCACCTCCGAGCAGCCACGAGGAGGAGG
CTTTCCTCTTTGAGCTTTTCGGGTCGTTCAAGTACTTCACGATGCCTGTTTGGATGGTC
TTGATCGTCACCTCGCTTACCTGGATCGGCTGGTTCCCTTTCATCCTCTTTGACACCGA
CTGGATGGGCCGAGAGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAAAAGT
ATCATGACGGTGTGAGAATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCTCGGG
ATCACATCTATTGGAATGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTATGGGG
TGTCTCCAATATCATCATGGCACTGTGCTTCGTGGCGATGCTTATTATAACATACGTTG
CGCAGAACCTGGATTATGGACCTAGTGGAGCACCTCCAACCGGCATTGTCGCCGCTTCC
CTCATTGTGTTCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTATACCATATGC
GATGGCTGCAAGTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGCATTC
TTAATTTATCTATTGTCATACCACAGATCATCGTGTCGCTG

Appendix 5 Nucleotide sequence similarity between the 5' end fragment of *TaSUT2* (query) and other *SUT2*s from the GenBank (subject).

EU693244.1 *Triticum aestivum* cv. Bobwhite *sucrose transporter 2*-like (*SUT2*) mRNA, complete sequence, length = 1843

Score = 1142 bits (618), Expect = 0.0
Identities = 628/633 (99%), Gaps = 0/633 (0%)
Strand = Plus/Plus

```
Query  26   GCGGGGGCACCTCCTCGGCCGCGCCGCCGCGCAAGGTCCCGCTCCGCTCGCTGCTCCGGG  85
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  180  GCGGGGGCACCTCCTCGGCCGCGCCGCCGCGCAAGGTCCCGCTCCGCTCGCTGCTCCGGG  239

Query  86   CGGCGTCCGTCGCCTGCGGGGTCCAATTCGGGTGGGCGCTGCAGCTCTCCCTGCTCACCC  145
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  240  CGGCGTCCGTCGCCTGCGGGGTCCAATTCGGGTGGGCGCTGCAGCTCTCCCTGCTCACCC  299

Query  146  CGTACGTCCAAGAACTCGGCATCCCGCACGCCTTCGCCAGCCTGGTCTGGCTCTGCGGCC  205
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  300  CGTACGTCCAAGAACTCGGCATCCCGCACGCCTTCGCCAGCCTGGTCTGGCTCTGCGGCC  359

Query  206  CGCTGTCGGGCCTCCTCGTGCAGCCCCTGGTGGGCCACCTCTCGGACCGCATCGCGCCGG  265
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  360  CGCTGTCGGGCCTCCTCGTGCAGCCCCTGGTGGGCCACCTCTCGGACCGCATCGCGCCGG  419

Query  266  CCAACTCCCCGCTGGGCCGCCGCCGGCCCTTCATCGCCGCGGGGGCCGCCTCCATCGCCT  325
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  420  CCAACTCCCCGCTGGGCCGCCGCCGGCCCTTCATCGCCGCGGGGGCCGCCTCCATCGCCT  479

Query  326  TCTCCGTGCTCACCGTCGGCTTCTCCGCCGACCTCGGCCGCCTCTTCGGGGACAACATCG  385
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Sbjct  480  TCTCCGTGCTCACCGTCGGCTTCTCCGCCGACCTCGGCCGCCTCTTCGGGGACAACACCG  539

Query  386  TGCCCGGGTCCACCCGGTTCGGCGCCATCATCGTCTACCTCATCGGCTTCTGGCTGCTGG  445
            |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  540  TGCCCGGGTCCGCCCGGTTCGGCGCCATCATCGTCTACCTCATCGGCTTCTGGCTGCTGG  599

Query  446  ATGTCGGCAACAACGCCACCCAGGGGCCATGCCGCGCCTTCCTCGCAGACCTCACAGAGA  505
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  600  ATGTCGGCAACAACGCCACCCAGGGGCCATGCCGCGCCTTCCTCGCAGACCTCACAGAGA  659

Query  506  ATGACCCGAGGAGGACCCGGATTGCCAACGCCTACTTCTCGCTCTTCATGGCCCTGGGGA  565
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  660  ATGACCCGAGGAGGACCCGGATTGCCAACGCCTACTTCTCGCTCTTCATGGCCCTGGGGA  719

Query  566  ACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGATATTCCCCTTCACTA  625
            |||||||||||||||||||||||||  |||||||||||||||| ||||||||||||||
Sbjct  720  ACATACTCGGGTACGCCACCGGGGCGCACAACGGCTGGTACAAGACATTCCCCTTCACTA  779

Query  626  TTACGGAGTCCTGCTCTGTCAGCTGCGCCAACC  658
            |||||||||||||||||||||||||||||||||
Sbjct  780  TTACGGAGTCCTGCTCTGTCAGCTGCGCCAACC  812
```

AJ272308.1 *Hordeum vulgare sucrose transporter 2* (*SUT2*) mRNA, complete sequence
length = 2018

Score = 861 bits (466), Expect = 0.0
Identities = 587/644 (91%), Gaps = 14/644 (2%)
Strand = Plus/Plus

```
Query  26   GCGGGGGCAC--C-TCCTCGGC--CG-CGCCGCCGCGCAAGGTCCCGCTCCGCTCGCTGC  79
            ||||||||||  | || |||||  || || ||||||||||||||| ||||| |||||||
Sbjct  162  GCGGGGGCACGTCGTCGTCGGCGGCGCCGGCGCCGCGCAAGGTCCCTCTCCGGTCGCTGC  221

Query  80   TCCGGGCGGCGTCCGTCGCCTGCGGGGTCCAATTCGGGTGGGCGCTGCAGCTCTCCCTGC  139
            |||||||||||| || || |||||||| || ||||||||||||||||||| |||||||
Sbjct  222  TCCGGGCGGCGTCGGTGGCGTGCGGGGTGCAGTTCGGGTGGGCGCTGCAGCTGTCCCTGC  281

Query  140  TCACCCCGTACGTCCAAGAACTCGGCATCCCGCACGCCTTCGCCAGCCTGGTCTGGCTCT  199
            | |||||||||||||| || || ||||||||||||||||||||||||||||| ||||| |
Sbjct  282  TGACCCCGTACGTCCAGGAGCTGGGCATCCCGCACGCCTTCGCCAGCCTGGTGTGGCTGT  341

Query  200  GCGGCCCGCTGT-CGGGCCTCCTCGTGCAGCCCCTGGTGGGCCACCTCTCGGACCGCATC  258
            |||||||||| | |||| ||||| ||||||||||||||||||||||||||||||||||||
Sbjct  342  GCGGCCCGCTCTCCGGG-CTCCTGGTGCAGCCCCTGGTGGGCCACCTCTCGGACCGCATC  400

Query  259  GCGCCGGCCAACTCCCCGCTGGGCCGCCGCCGGCCCTTCATCGCCGCGGGGGCCGCCTCC  318
             |||||||||||||||||| |||||||||||||||||||||||| ||||| ||||||
Sbjct  401  ACGCCGGCCAACTCCCCGCTCGGCCGCCGCCGGCCCTTCATCGCCGCCGGGGCGGCCTCC  460

Query  319  ATCGCCTTCTCCGTGCTCACCGTCGGCTTCTCCGCCGACCTCGGCCGCCTCTTCGGGGAC  378
            |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  461  ATCGCCTTCGCCGTGCTCACCGTCGGCTTCTCCGCCGACCTCGGCCGCCTCTTCGGGGAC  520

Query  379  AACATCGTGCCCGGGTCCACCCGGTTCGGCGCCATCATCGTCTACCTCATCGGCTTCTGG  438
            ||| | ||||| || ||||| ||  ||||||||||||||||||||| ||||||||||||
Sbjct  521  AACGTGGTGCCTGGCTCCACACGCATCGGCGCCATCATCGTCTACCTCGTCGGCTTCTGG  580

Query  439  CTGCTGGATGTCGGCAACAACGCCACCCAGGGGCCATGCCGCGCCTTCCTCGCAGACCTC  498
            ||||| || ||||||||||||||||||||| |||||||||||||||||||| ||||||
Sbjct  581  CTGCTCGACGTCGGCAACAACGCCACCCAGGGCCCATGCCGCGCCTTCCTCGCCGACCTC  640

Query  499  ACAGAGAATGACCCGAGGAGGACCCGGATTGCCAACGCCTACTTCTCGCTCTTCATGGCC  558
            || |||||||||||||||||||||||| ||||| |||||||||||| ||||||||||||
Sbjct  641  ACCGAGAATGACCCGAGGAGGACCCGGATCGCCAATGCCTACTTCTCACTCTTCATGGCC  700

Query  559  CTGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGATATTCCCC  618
            |||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||
Sbjct  701  CTGGGGAACATACTCGGGTACGCCACCGGGGCGTACAATGGCTGGTACAAGATATTCCCG  760

Query  619  TTCACTATTAC-GGAGTCCTGCT--CTGTCAGCTGCGCCAACCT  659
            |||||||| || || |||||| |  | ||||||||||||||||
Sbjct  761  TTCACTATCACTGG-GTCCTG-TGGC-GTCAGCTGCGCCAACCT  801
```

GU045300.1 *Sorghum bicolor sucrose transporter 4* (*SUT4*) mRNA, complete sequence, length = 1835

Score = 608 bits (329), Expect = 1e-170
Identities = 522/616 (85%), Gaps = 10/616 (2%)
Strand = Plus/Plus

```
Query  49   CCGCCGCGCAAGGT-CCCGCTCCGCTCGCTGCTCCGGGCGGCGTCCGTCGCCTGCGGGGT  107
            |||||||| ||||| ||| |||||   |||||| || |||||||| ||||||||||||||
Sbjct  73   CCGCCGCGGAAGGTGCCC-CTCCGGAAGCTGCTGCGTGCGGCGTCGGTCGCCTGCGGGGT  131

Query  108  CCAATTCGGGTGGGCGCTGCAGCTCTCCCTGCTCACCCCGTACGTCCAAGAACTCGGCAT  167
             || ||||| ||||||||||||| ||  |||| |||||||||| || || || |||||
Sbjct  132  GCAGTTCGGCTGGGCGCTGCAGCTGTCGTTGCTGACCCCGTACGTGCAGGAGCTGGGCAT  191

Query  168  CCCGCACGCCTTCGCCAGCCTGGTCTGGCTCTGCGGCCCGCTGTCGGGCCTCCTCGTGCA  227
            |||||||||||||||||||| |||||||| |||||||||||||||||||||||||||| ||
Sbjct  192  CCCGCACGCCTTCGCCAGCCTCGTCTGGCTGTGCGGCCCGCTGTCGGGCCTCCTCGTCCA  251

Query  228  GCCCCTGGTGGGCCACCTCTCGGACCGCATCGCGCCGGCCAACTCCCCGCT-GGGCCGCC  286
            |||||| || |||||||||| |||||| |||| || |||  ||| ||||| |||| |||
Sbjct  252  GCCCCTCGTCGGCCACCTCTCCGACCGCCTCGCCCCCGCCGCCTCGCCGCTCGGGC-GCC  310

Query  287  GCCGGCCCTTCATCGCCGCGGGGGCCGCCTCCATCGCCTTCT-CCGTGCTCACCGTCGGC  345
            || |||||||||||||| || ||||| |||||||||  |  ||||||||||||||||||
Sbjct  311  GCAGGCCCTTCATCGCCGCCGGCGCCGCGTCCATCGCCG-CAGCCGTGCTCACCGTCGGC  369

Query  346  TTCTCCGCCGACCTCGGCCGCCTCTTCGGGGACAACATCGTGCCCGGG-TCCACCCGGTT  404
            |||||||||||||||||| |||||||| ||| || ||   |||||| || || ||  |
Sbjct  370  TTCTCCGCCGACCTCGGCCGACTCTTCGGCGACGACGTCAC-CCCGGGCTCAACGCGCCT  428

Query  405  CGGCGCCATCAT-CGTCTACCTCATCGGCTTCTGGCTGCTGGATGTCGGCAACAACGCCA  463
            |||||||||| | |||||||||| |||| |||||||||| || ||||||||||||||||
Sbjct  429  CGGCGCCATC-TGCGTCTACCTCGTCGGATTCTGGCTGCTCGACGTCGGCAACAACGCCA  487

Query  464  CCCAGGGGCCATGCCGCGCCTTCCTCGCAGACCTCACAGAGAATGACCCGAGGAGGACCC  523
            | |||||||| ||| | |||||||||| |||||||||||||| ||||| |||||||| |
Sbjct  488  CGCAGGGGCCCTGCAGGGCCTTCCTCGCCGACCTCACAGAGAACGACCCAAGGAGGACTC  547

Query  524  GGATTGCCAACGCCTACTTCTCGCTCTTCATGGCCCTGGGGAACATACTCGGGTACGCCA  583
            |||| || || || ||||| || |||||||||||||||| |||||||| || || ||||
Sbjct  548  GGATCGCTAATGCATACTTTTCACTCTTCATGGCCCTGGGAAACATACTTGGATATGCCA  607

Query  584  CCGGGGCATACAACGGCTGGTACAAGATATTCCCCTTCACTATTACGGAGTCCTGCTCTG  643
            |||| |||||||  || |||||   |||||| || |||||| |||| |||||||||   |
Sbjct  608  CCGGAGCATACAGTGGATGGTATTCGATATTTCCTTTCACTGTTACAGAGTCCTGCGGAG  667

Query  644  TCAGCTGCGCCAACCT  659
            |||| || ||||||||
Sbjct  668  TCAGTTGTGCCAACCT  683
```

GQ485583.1 *Saccharum hybrid* cv. ROC22 *sucrose transporter 4* (*SUT4*) mRNA, complete sequence, length = 1861

Score = 582 bits (315), Expect = 9e-163
Identities = 522/621 (84%), Gaps = 17/621 (3%)
Strand = Plus/Plus

```
Query  49   CCGCCGCGCAAGGT-CCCGCTCCGCTCGCTGCTCCGGGCGGCGTCCGTCGCCTGCGGGGT  107
            |||||||| ||||| ||| |||||   |||||| || |||||||| ||||||||||||||
Sbjct  83   CCGCCGCGGAAGGTGCCC-CTCCGGAAGCTGCTGCGTGCGGCGTCGGTCGCCTGCGGGGT  141

Query  108  CCAATTCGGGTGGGCGCTGCAGCTCTCCCTGCTCACCCCGTACGTCCAAGAACTCGGCAT  167
             || ||||| ||||||||||||||| || ||||| |||||||||||| || || || |||||
Sbjct  142  GCAGTTCGGCTGGGCGCTGCAGCTGTCGCTGCTGACCCCGTACGTGCAGGAGCTGGGCAT  201

Query  168  CCCGCACGCCTTCGCCAGCCTGGTCTGGCTCTGCGGCCCGCTGTCGGGCCTCCTCGTGCA  227
            |||||||||||||||||||| |||||||| |||||||||||||||||||||||||||| ||
Sbjct  202  CCCGCACGCCTTCGCCAGCCTCGTCTGGCTGTGCGGCCCGCTGTCGGGCCTCCTCGTTCA  261

Query  228  GCCCCTGGTGGGCCACCTCTCGGACCGCATC-GCGCCGGCCAACTCCCCGCT-GGGCCGC  285
            |||||| || |||||||||| ||||||||| || || |||  ||| ||||| ||| |||
Sbjct  262  GCCCCTCGTCGGCCACCTCTCCGACCGCATCGGC-CCCGCCGCCTCGCCGCTCGGG-CGC  319

Query  286  CGCCGGCCCTTCATCGCCGCGGGGGCCGCCTCCATCGCCTTC-TCCGTGCTCACCGTCGG  344
            ||| ||||||||||||||| || ||||| ||||||| |  |  ||||||||||||||||
Sbjct  320  CGCAGGCCCTTCATCGCCGCCGGCGCCGCGTCCATCG-CAGCAGCCGTGCTCACCGTCGG  378

Query  345  CTTCTCCGCCGACCTCGGCCGCCTCTT-CGGGGACAACATCGTGCCCGGG-TCCACCCGG  402
            |||||||||||||||||||| ||||| ||| ||| || ||   |||||| || || ||
Sbjct  379  CTTCTCCGCCGACCTCGGCCGACTCTTGCGGCGACGACGTCAC-CCCGGGCTCAACGCGC  437

Query  403  TTCGGCGCCATCAT-CGTCTACCTCATCGGCTTCTGGCTGCTGGATGTCGGCAACAACGC  461
             |||||||||| | |||||||||| | || |||||||||| || ||||||||||||||
Sbjct  438  CTCGGCGCCATC-TGCGTCTACCTCGTAGGATTCTGGCTGCTCGACGTCGGCAACAACGC  496

Query  462  CACCCAGGGGCCATGCCGCGCC-TTC-CTCGCAGACCTCACAGAGAATGACCCGAGGAGG  519
            ||| |||||||| ||| | | | ||| ||||| |||||||||||||| ||||| ||||||
Sbjct  497  CACGCAGGGGCCCTGCGGGGGCGTTCGCTCGCCGACCTCACAGAGAACGACCCAAGGAGG  556

Query  520  ACCCGGATTGCCAACGCCTACTTCTCGCTCTTCATGGCCCTGGGGAACATACTCGGGTAC  579
            || ||||| || || || ||||| || |||||||||||||||| |||||||| || ||
Sbjct  557  ACTCGGATCGCTAATGCATACTTTTCACTCTTCATGGCCCTGGGAAACATACTTGGATAT  616

Query  580  GCCACCGGGGCATACAACGGCTGGTACAAGATATTCCCCTTCACTATTACGGAGTCCTGC  639
            |||||||| |||||||  || |||||   |||||| || |||||| |||| |||||||||
Sbjct  617  GCCACCGGAGCATACAGTGGATGGTATTTGATATTTCCTTTCACTGTTACAGAGTCCTGC  676

Query  640  T-CTGTCAGCTGCGCCAACCT  659
              | ||||| || ||||||||
Sbjct  677  GGC-GTCAGTTGTGCCAACCT  696
```

DQ072592.1 *Oryza sativa* Indica Group *sucrose transporter 2* (*SUT2M)* mRNA, complete sequence, length = 1531

Score = 582 bits (315), Expect = 9e-163
Identities = 488/573 (85%), Gaps = 6/573 (1%)
Strand = Plus/Plus

```
Query  54   GCGCAAGGTCCCGCTCCGCTCGCTGCTCCGGGCGGCGTCCGTCGCCTGCGGGGTCCAATT  113
            ||||||||||||  ||  ||   |||  ||  ||  ||||||||  ||  ||  |||||  ||  ||  ||
Sbjct  78   GCGCAAGGTCCCCCTGCGGAAGCTTCTGCGTGCGGCGTCGGTGGCGTGCGGCGTGCAGTT  137

Query  114  CGGGTGGGCGCTGCAGCTCTCCCTGCTCACCCCGTACGTCCAAGAACTCGGCATCCCGCA  173
            ||||||||||||||||||||||||||||  |||||  ||||||||  ||  ||||||||||||||
Sbjct  138  CGGGTGGGCGCTGCAGCTCTCCCTGCTGACCCCCTACGTCCAGGAGCTCGGCATCCCGCA  197

Query  174  CGCCTTCGCCAGCCTGGTCTGGCTCTGCGGCCCGCTGTCGGGCCTCCTCGTGCAGCCCCT  233
            ||||||||||||||  ||||||||||||||||||  ||  ||  ||||||||||  ||||||||
Sbjct  198  CGCCTTCGCCAGCCTCGTCTGGCTCTGCGGCCCCCTCTCCGGCCTCCTCGTCCAGCCCCT  257

Query  234  GGTGGGCCACCTCTCGGACCGCATCGCGCCGGCCAACTCCCCGCTGGGCCGCCGCCGGCC  293
             ||  ||||||||||  ||||||||||  ||  |||  ||||||  ||  ||||||||||  ||
Sbjct  258  CGTCGGCCACCTCTCCGACCGCATCGCCCCCGCCGCCTCCCCCCTCGGCCGCCGCCGCCC  317

Query  294  CTTCATCGCCGCGGGGGCCGCCTCCATCGCCTTCTC-C-GTGCTCACCGTCGGCTTCTCC  351
            ||||||||||||  ||  |||||  ||||||||  |  |  |  |  ||  ||||||||||  ||||||
Sbjct  318  CTTCATCGCCGCCGGCGCCGCTTCCATCGC-TGC-CGCAGTCCTCACCGTCGGATTCTCC  375

Query  352  GCCGACCTCGGCCGCCTCTTCGGGGACAACATCGTGCCCGGG-TCCACCCGGTTCGGCGC  410
            ||||||||||||||  |||||||  ||  ||||  ||||||  ||||||||  |||||||
Sbjct  376  GCCGACCTCGGCCGAATCTTCGGCGATTCCATCAC-CCCGGGCTCCACCCGCCTCGGCGC  434

Query  411  CATCATCGTCTACCTCATCGGCTTCTGGCTGCTGGATGTCGGCAACAACGCCACCCAGGG  470
            ||||||||||||||||  ||||||||||||||  ||  ||  ||||||||||||||  ||  |||||
Sbjct  435  CATCATCGTCTACCTCGTCGGCTTCTGGCTCCTCGACGTCGGCAACAACGCTACACAGGG  494

Query  471  GCCATGCCGCGCCTTCCTCGCAGACCTCACAGAGAATGACCCGAGGAGGACCCGGATTGC  530
             ||  |||  |  ||||||||||  ||||||||  ||||||||||  ||||||||  |||||  ||
Sbjct  495  ACCCTGCAGGGCCTTCCTCGCCGACCTCACCGAGAATGACCCAAGGAGGACTCGGATAGC  554

Query  531  CAACGCCTACTTCTCGCTCTTCATGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGC  590
             ||  ||  ||||||||  |  ||||||||||||||  ||||||||  ||  ||  |||||  ||  ||
Sbjct  555  TAATGCTTACTTCTCATTGTTCATGGCCCTGGGAAACATACTTGGATATGCCACTGGAGC  614

Query  591  ATACAACGGCTGGTACAAGATATTCCCCTTCAC  623
            |||||  ||||||||||||||||||||  |||||
Sbjct  615  ATACAGTGGCTGGTACAAGATATTCCCGTTCAC  647
```

Appendix 6 Nucleotide sequence similarity between the 3’ end fragment of *TaSUT2* and other *SUT2*s from the GenBank.

EU693244.1 *Triticum aestivum* cv. Bobwhite *sucrose transporter 2*-like (*SUT2*) mRNA, complete sequence, length = 1843

Score = 893 bits (483), Expect = 0.0
Identities = 493/498 (99%), Gaps = 0/498 (0%)
Strand = Plus/Plus

```
Query  1     CGCCGCTTCCCTCATTGTGTTCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTAT  60
             |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Sbjct  1346  CGCCGCTTCCCTCATTGTGTTCACAATCCTAGGAGCACCCCTGTCGGTTACGTACAGTAT  1405

Query  61    ACCATATGCGATGGCTGCAAGTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAAT  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1406  ACCATATGCGATGGCTGCAAGTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAAT  1465

Query  121   GGGCATTCTTAATTTATCTATTGTCATACCACAGATCATCGTGTCGCTGGGCAGCGGGCC  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1466  GGGCATTCTTAATTTATCTATTGTCATACCACAGATCATCGTGTCGCTGGGCAGCGGGCC  1525

Query  181   ATGGGACCAGCTCTTCGGCGGAGGGAACGCGCCATCCTTCTGGGTGGCTGCCGCAGCCTC  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1526  ATGGGACCAGCTCTTCGGCGGAGGGAACGCGCCATCCTTCTGGGTGGCTGCCGCAGCCTC  1585

Query  241   CTTTGTGGGCGGGCTGGTAGCCATCCTGGGGCTCCCGCGGGCCCGGCTCGGGCCGAAGAA  300
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct  1586  CTTTGTGGGCGGGCTGGTAGCCATCCTGGGGCTCCCGCGGGCCCGGCTTGGGCCGAAGAA  1645

Query  301   GAAAACCACCCAACGATGATGATTCGAGCGAGTATACAGAGACATTGTTATGTTGCAGAC  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1646  GAAAACCACCCAACGATGATGATTCGAGCGAGTATACAGAGACATTGTTATGTTGCAGAC  1705

Query  361   AAGGAAGGAAGGCAGGTTGTACACTGGTTCATTCTTTGCACACTGGTTCCAAGACACCGG  420
             |||||||||||| |||||||||||||||| |||||||||||||||||||||||||||||
Sbjct  1706  AAGGAAGGAAGGTAGGTTGTACACTGGTTCACTCTTTGCACACTGGTTCCAAGACACCGG  1765

Query  421   GAGCGCTGCCTACACTACACTAGGCATGGGGGGTCTGTATATTGTCATCATCAAGTTGGA  480
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1766  GAGCGCTGCCGACACTACACTAGGCATGGGGGGTCTGTATATTGTCATCATCAAGTTGGA  1825

Query  481   TTGGGTTCAAATTCAAGG  498
             ||||||||||||||||||
Sbjct  1826  TTGGGTTCAAATTCAAGG  1843
```

AJ272308.1 *Hordeum vulgare sucrose transporter 2 (SUT2)* mRNA, complete sequence, length = 2018

Score = 854 bits (462), Expect = 0.0
Identities = 618/686 (90%), Gaps = 40/686 (6%)
Strand = Plus/Plus

```
Query  1     CGCCGCTTCCCTCATTGTGTTCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTAT  60
             |||||||||||||||||| ||||||||| |||||||||||||||| | ||||||||||
Sbjct  1333  CGCCGCTTCCCTCATTGTTTTCACAATCTTAGGAGCACCTCTGTCGATCACGTACAGTAT  1392

Query  61    ACCATATGCGATGGCTGCAAGTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAAT  120
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1393  ACCGTATGCGATGGCTGCAAGTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAAT  1452

Query  121   GGGCATTCTTAATTTATCTATTGTCATACCACAGATCATCGTGTCGCTGGGCAGCGGGCC  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1453  GGGCATTCTTAATTTATCTATTGTCATACCACAGATCATCGTGTCGCTGGGCAGCGGGCC  1512

Query  181   ATGGGACCAGCTCTTCGGCGGAGGGAACGCGCCATCCTTCTGGGTGGCTGCCGCAGCCTC  240
              |||||||||||||||||||||||||||||| |||||||||||||||||| |||||
Sbjct  1513  GTGGGACCAGCTCTTCGGCGGAGGGAACGCGCCGTCCTTCTGGGTGGCTGCCGCGGCCTC  1572

Query  241   CTTTGTGGGCGGGCTGGTAGCCATCCTGGGGCTCCCGCGGGCCCGGCTCGGGCCGAAGAA  300
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1573  ATTCGTGGGCGGGCTGGTAGCCATCCTGGGGCTCCCGCGGGCCCGGCTCGGGCCGAAGAA  1632

Query  301   GAAAACCACCCAACGATGATGAT--TC--GA---G--CGAGTATACAGAGACATTGTTAT  351
             ||||||||||||||||||||||  |   ||   |  |||||||| ||||||||||||||
Sbjct  1633  GAAAACCACCCAACGATGATGATGATGATTTGAGCGAGTATAGAGAGACATTGTTAT  1692

Query  352   GTTGCAGACAAGGAAGGAAGGCAGGTTGTACAC---TGGT-TCA--TT-C-----TTT--  397
             |||||||||||||||||||||||||||||||   |||| |||  || |     |||
Sbjct  1693  GTTGCAGACAAGGAAGGAAGGCAGGTTGTACACACATGGTCTCAAGTTGCAACCATTTAT  1752

Query  398   -----GCACACTGGTTCCAAGACACCGGGAGCGCTGCCTACACTACACTAGGCATGGGGG  452
                  |||||| ||||||  | |||||||| |||||||||||| ||||||||||||||||
Sbjct  1753  TCTTTGCACACCGGTTCCGTGGCACCGGGAACGCTGCCTACACCACACTAGGCATGGGGG  1812

Query  453   G-TCTGTATATTGTCATCAT---CA---AGTTGGATTGGGTTCAAATTCAAGGCCAAAAG  505
             | ||||||||||||||||   ||   |||||| |||||||||||||||| ||||||
Sbjct  1813  GGTCTGTATATTGTCATCATTATCATCGAGTTGGGTTGGGTTCAAATTCAAGGGCAAAAG  1872

Query  506   TAGGAGGGCTCTGGAATTTTTCTTCCTCTTGTTACTTTCCTCCATGGGCCCTATCATTGC  565
             |||||||| ||||| |||||||||||||||||||||||| |||||||||| | ||
Sbjct  1873  TAGGAGGGCCCTGGAGTTTTTCTTCCTCTTGTTACTTTCCTCCGTGGGCCCTATCGTCGC  1932

Query  566   Ctttttt--C-GCTTGTCGGTGTGAAACATGGTCAGTCATATGTTGTTGTA-TAT-TCG  620
             |||||||  | |||||||||||| |||||||||||||||||||||||  ||| |||
Sbjct  1933  CTTTTTTTTTCCGCTTGTCGGTGTGGAACATGGTCAGTCATATGTTGTTGTTGTATATCG  1992

Query  621   GATATTTTGTCATTTTGTCGGAGATT  646
             ||| ||||||||||||||||||||||
Sbjct  1993  GATGTTTTGTCATTTTGTCGGAGATT  2018
```

XM_003577230.1 PREDICTED: *Brachypodium distachyon sucrose transporter 2*-like (*SUT2*-like) mRNA, complete sequence, length = 1925

Score = 292 bits (158), Expect = 2e-75
Identities = 241/282 (85%), Gaps = 2/282 (1%)
Strand = Plus/Plus

```
Query  5     GCTTCCCTCATTGTGTTCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTATACCA  64
             ||||||  |  ||  ||  ||  |||||   ||||||||||  |||  ||||  ||  ||||||||||||
Sbjct  1241  GCTTCCATTATAGTTTTTACAATTTTAGGAGCACCACTGGCGGTCACATACAGTATACCA  1300

Query  65    TATGCGATGGCTGCAAGTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGC  124
             ||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Sbjct  1301  TATGCGATGGCTGCAAGTCGTGTTGAAAATCTTGGGCTTGGCCAAGGTCTAGCAATGGGC  1360

Query  125   ATTCTTAATTTATCTATTGTCATACCACAGATCATCGTGTCGCTGGGCAGCGGGCCATGG  184
             |||||  ||||||||||||||||  ||||||||||  ||||||||||  ||  |||||  |||
Sbjct  1361  ATTCTCAATTTATCTATTGTCATTCCACAGATCATTGTGTCGCTGGGTAGTGGGCCGTGG  1420

Query  185   GACCAGCTCTTCGGCGGAGGGAACGCGCCATCCTTCTGGGTGGCTGCCGCAGCCTCCTTT  244
             |||||  ||||||||||  ||  ||  |||||   |  ||  |   |||||  ||  ||  |||||  |||
Sbjct  1421  GACCAACTCTTCGGCGGTGGAAATGCGCCGGCATTTTTTGTGGCAGCTGCCGCCTC-TTT  1479

Query  245   -GTGGGCGGGCTGGTAGCCATCCTGGGGCTCCCGCGGGCCCG  285
              ||  ||  ||||||||  ||  ||||||||  ||  ||  ||  |||||
Sbjct  1480  CGTCGGTGGGCTGGTTGCTATCCTGGGACTTCCACGAGCCCG  1521
```

DQ072592.1 *Oryza sativa* Indica Group *sucrose transporter 2* (*SUT2M)* mRNA, complete sequence, length = 1531

Score = 231 bits (125), Expect = 4e-57
Identities = 231/283 (82%), Gaps = 4/283 (1%)
Strand = Plus/Plus

```
Query  5     GCTTCCCTCATTGTGTTCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTATACCA  64
             ||||||||  | || || |||||  ||||||| || ||| || | ||||||||||||||
Sbjct  1216  GCTTCCCTGGTAGTTTTTACAATTTTAGGAGCGCCCCTGGCGATCACGTACAGTATACCA  1275

Query  65    TATGCGATGGCTGCAAGTCGTGTTGAAAATCTTGGG-CTAGGCCAAGGTCTAGCAATGGG  123
             ||||| |||||||| ||||| |||||||||| |||| || ||||||||||||||||||||
Sbjct  1276  TATGCAATGGCTGCTAGTCGGGTTGAAAATC-TGGGACTTGGCCAAGGTCTAGCAATGGG  1334

Query  124   CATTCTTAATTTATCTATTGTCATACCACAGATCATCGTGTCGCTGGGCAGCGGGCCATG  183
             ||||||||||||  |||||||||||||||| | || ||||| ||||| |||||||| ||
Sbjct  1335  CATTCTTAATTTGGCTATTGTCATACCACAGGTTATTGTGTCACTGGGTAGCGGGCCCTG  1394

Query  184   GGACCAGCTCTTCGGCGGAGGGAACGCGCCATCCTTCTGG-GTGGCTGCCGCAGCCTCCT  242
             |||||| || || || || || || || ||| |||| ||  |||||||| || || || |
Sbjct  1395  GGACCAACTGTTTGGTGGTGGCAATGCACCAGCCTT-TGCAGTGGCTGCTGCTGCATCTT  1453

Query  243   TTGTGGGCGGGCTGGTAGCCATCCTGGGGCTCCCGCGGGCCCG  285
             || | || |||||||| || || ||||| || || || |||||
Sbjct  1454  TTATCGGTGGGCTGGTGGCTATTCTGGGCCTTCCACGAGCCCG  1496
```

AB091672.1 *Oryza sativa* Japonica Group *sucrose transporter 2 (SUT2)* mRNA, complete sequence, length = 1869

Score = 231 bits (125), Expect = 4e-57
Identities = 231/283 (82%), Gaps = 4/283 (1%)
Strand = Plus/Plus

```
Query  5     GCTTCCCTCATTGTGTTCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTATACCA  64
             ||||||||  | || || |||||  ||||||| || ||| || | ||||||||||||||||
Sbjct  1270  GCTTCCCTGGTAGTTTTTACAATTTTAGGAGCGCCCCTGGCGATCACGTACAGTATACCA  1329

Query  65    TATGCGATGGCTGCAAGTCGTGTTGAAAATCTTGGG-CTAGGCCAAGGTCTAGCAATGGG  123
             ||||| |||||||| ||||| |||||||||| |||| || ||||||||||||||||||||
Sbjct  1330  TATGCAATGGCTGCTAGTCGGGTTGAAAATC-TGGGACTTGGCCAAGGTCTAGCAATGGG  1388

Query  124   CATTCTTAATTTATCTATTGTCATACCACAGATCATCGTGTCGCTGGGCAGCGGGCCATG  183
             ||||||||||||  |||||||||||||||| | || ||||| ||||| |||||||| ||
Sbjct  1389  CATTCTTAATTTGGCTATTGTCATACCACAGGTTATTGTGTCACTGGGTAGCGGGCCCTG  1448

Query  184   GGACCAGCTCTTCGGCGGAGGGAACGCGCCATCCTTCTGG-GTGGCTGCCGCAGCCTCCT  242
             |||||| || || || || || || || ||| |||| ||  |||||||| || || || |
Sbjct  1449  GGACCAACTGTTTGGTGGTGGCAATGCACCAGCCTT-TGCAGTGGCTGCTGCTGCATCTT  1507

Query  243   TTGTGGGCGGGCTGGTAGCCATCCTGGGGCTCCCGCGGGCCCG  285
             || | || |||||||| || || ||||| || || || |||||
Sbjct  1508  TTATCGGTGGGCTGGTGGCTATTCTGGGCCTTCCACGAGCCCG  1550
```

Appendix 7. Nucleotide sequence of the initially isolated 5'end fragment of *TaSUT2*. The nested primer sets used for extending the 5'end are underlined.

ACGCGGGGGCACCTCCTCGGCCGCGCCGCCGCGCAAGGTCCCGCTCCGCTCGCTGCTCC
GGGCGGCGTCCGTCGCCTGCGGGGTCCAATTCGGGTGGGCGCTGCAGCTCTCCCTGCTC
ACCCCGTACGTCCAAGAACTCGGCATCCCGCACGCCTTCGCCAGCCTGGTCTGGCTCTG
CGGCCCGCTGTCGGGCCTCCTCGTGCAGCCCCTGGTGGGCCACCTCTCGGACCGCATCG
CGCCGGCCAACTCCCCGCTGGGCCGCCGCCGGCCCTTCATCGCCGCGGGGGCCGCCTCC
ATCGCCTTCTCCGTGCTCACCGTCGGCTTCTCCGCCGACCTCGGCCGCCTCTTCGGGGA
CAACATCGTGCCCGGGTCCACCCGGTTCGGCGCCATCATCGTCTACCTCATCGGCTTCT
GGCTGCTGGATGTCGGCAACAACGCCACCCAGGGGCCATGCCGCGCCTTCCTCGCAGAC
CTCACAGAGAATGACCCGAGGAGGACCCGGATTGCCAACGCCTACTTCTCGCTCTTCAT
GGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTACAAGATAT
TCCCCTTCACTATTACGGAGTCCTGCTCTGTCAGCTGCGCCAACCT

Appendix 8. Nucleotide sequence of the full length DNA fragment of *TaSUT2*. Sequences of the forward and reverse primers used for amplifying the full length sequence of *TaSUT2* are underlined.

ACGCGGGGAGAGCGATAGATACTTTGCTTTCCCCTCCCTCCATCCCTTCCCTTCCCTCT
CTCCCAGGTCAGGCCAGGTCTCCCACAGAATTCAGCCAGGAGGCGTGGGGGGAGCGGGC
AGCCAAGAAAGTCCAACACCACTCATCAAGGGCGATGCCGCCGCGGCGGCCCAACACCG
GCGCTGGGGGCGGGGGCACCTCCTCGGCCGCGCCGCCGCGCAAGGTCCCGCTCCGCTCG
CTGCTCCGGGCGGCGTCCGTCGCCTGCGGGGTCCAATTCGGGTGGGCGCTGCAGCTCTC
CCTGCTCACCCCGTACGTCCAAGAACTCGGCATCCCGCACGCCTTCGCCAGCCTGGTCT
GGCTCTGCGGCCCGCTGTCGGGCCTCCTCGTGCAGCCCCTGGTGGGCCACCTCTCGGAC
CGCATCGCGCCGGCCAACTCCCCGCTGGGCCGCCGCCGGCCCTTCATCGCCGCGGGGGC
CGCCTCCATCGCCTTCTCCGTGCTCACCGTCGGCTTCTCCGCCGACCTCGGCCGCCTCT
TCGGGGACAACATCGTGCCCGGGTCCACCCGGTTCGGCGCCATCATCGTCTACCTCATC
GGCTTCTGGCTGCTGGATGTCGGCAACAACGCCACCCAGGGGCCATGCCGCGCCTTCCT
CGCAGACCTCACAGAGAATGACCCGAGGAGGACCCGGATTGCCAACGCCTACTTCTCGC
TCTTCATGGCCCTGGGGAACATACTCGGGTACGCCACCGGGGCATACAACGGCTGGTAC
AAGATATTCCCCTTCACTATTACGGAGTCCTGCTCTGTCAGCTGCGCCAACCTCAAGTC
CGCGTTCCTGCTCGACATCATCATCCTGGCGATTACGACGTACGTTAGCGTCGTGACGG
TGCAGGACAATCCTACTTTCGGAAGCGACGAGGCGGCACCTCCGAGCAGCCACGAGGAG
GAGGCTTTCCTCTTTGAGCTTTTCGGGTCGTTCAAGTACTTCACGATGCCTGTTTGGAT
GGTCTTGATCGTCACCTCGCTTACCTGGATCGGCTGGTTCCCTTTCATCCTCTTTGACA
CCGACTGGATGGGCCGAGAGATCTACCGGGGAAGCCCGGAGATCGTCGCCGACACCCAA
AAGTATCATGACGGTGTGAGAATGGGCTCTTTTGGTCTCATGCTCAACTCGGTTGTTCT
CGGGATCACATCTATTGGAATGGAGAAGTTGTGTAGGAAGTGGGGAGCTGGACTTGTAT
GGGGTGTCTCCAATATCATCATGGCACTGTGCTTCGTGGCGATGCTTATTATAACATAC
GTTGCGCAGAACCTGGATTATGGACCTAGTGGAGCACCTCCAACCGGCATTGTCGCCGC
TTCCCTCATTGTGTTCACAATCCTAGGAGCACCTCTGTCGGTTACGTACAGTATACCAT
ATGCGATGGCTGCAAGTCGTGTTGAAAATCTTGGGCTAGGCCAAGGTCTAGCAATGGGC
ATTCTTAATTTATCTATTGTCATACCACAGATCATCGTGTCGCTGGGCAGCGGGCCATG
GGACCAGCTCTTCGGCGGAGGGAACGCGCCATCCTTCTGGGTGGCTGCCGCAGCCTCCT
TTGTGGGCGGGCTGGTAGCCATCCTGGGGCTCCCGCGGGCCCGGCTCGGGCCGAAGAAG
AAAACCACCCAACGATGATGATTCGAGCGAGTATACAGAGACATTGTTATGTTGCAGAC
AAGGAAGGAAGGCAGGTTGTACACTGGTTCATTCTTTGCACACTGGTTCCAAGACACCG
GGAGCGCTGCCTACACTACACTAGGCATGGGGGGTCTGTATATTGTCATCATCAAGTTG
GATTGGGTTCAAATTCAAGGCCAAAAGTAGGAGGGCTCTGGAATTTTTCTTCCTCTTGT
TACTTTCCTCCATGGGCCCTATCATTGCCTTTTTTTCGCTTGTCGGTGTGAAACATGGT
CAGTCATATGTTGTTGTATATTCGGATATTTTGTCATTTTGTCGGAGATTATCAATCAA
TCAATAAAAGTTGCAGGTTTTTGTTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Appendix 9. Analysis of variance for relative abundance of *TaSUT2* transcript in developing wheat grains

Time	Estimate	Error	Group
25DAA	3.3659	0.3134	A
5DAA	2.8500	0.2563	AB
10DAA	2.2659	0.3134	ABC
15DAA	1.6100	0.2563	BC
30DAA	1.4600	0.2563	C
20 DAA	1.0633	0.2563	C

Different letters indicate significant difference among developmental stages using Tukey's LSD, $p<0.05$.

Appendix 10. Analysis of variance for relative abundance of *TaSUT2* transcript in flag leaf blade during grain development

Time	Estimate	Error	Group
10DAA	19.4550	1.8147	A
-8DAA	14.8667	1.4817	AB
5DAA	13.2450	1.8147	ABC
25DAA	7.8350	1.8147	BC
15DAA	6.4400	1.8147	BC
20DAA	5.4300	1.8147	C
30 DAA	4.1233	1.4817	C

Different letters indicate significant difference among developmental stages using Tukey's LSD, $p<0.05$.

Appendix 11. Analysis of variance for relative abundance of *TaSUT2* transcript in flag leaf sheath during grain development

Time	Estimate	Error	Group
25DAA	7.3600	0.5199	A
10DAA	7.2900	0.4245	A
5DAA	5.1350	0.5199	AB
-8DAA	4.8267	0.4245	B
30DAA	4.3867	0.4245	B
20DAA	4.1067	0.4245	B
15 DAA	3.7733	0.4245	B

Different letters indicate significant difference among developmental stages usingTukey's LSD, $p<0.05$.

Appendix 12. Analysis of variance for relative abundance of *TaSUT2* transcript in peduncle during grain development

Time	Estimate	Error	Group
15DAA	3.9300	1.1569	A
30DAA	3.7967	0.9446	A
25DAA	3.3567	0.9446	A
20DAA	3.2900	1.1569	A
10DAA	2.7300	1.1569	A
-8DAA	2.3467	0.9446	A
5 DAA	1.5767	0.9446	A

Different letters indicate significant difference among developmental stages using Tukey's LSD, $p<0.05$.

Appendix 13. Analysis of variance for relative abundance of *TaSUT2* transcript in glume during grain development

Time	Estimate	Error	Group
30DAA	2.7767	0.3264	A
10DAA	2.6300	0.3264	A
15DAA	2.3200	0.3264	A
25DAA	2.2300	0.3264	A
20DAA	1.8867	0.3264	A
5DAA	1.5900	0.3264	A

Different letters indicate significant difference among developmental stages using Tukey's LSD, $p<0.05$.

Appendix 14. Analysis of variance for relative abundance of *TaSUT2* transcript in lemma during grain development.

Statistical analysis of relative abundance of *TaSUT2* transcript in lemma			
Time	Estimate	Error	Group
30DAA	5.2933	0.7004	A
5DAA	4.8272	0.8576	A
10DAA	4.6067	0.7004	A
25DAA	4.5933	0.7004	A
15DAA	3.9267	0.7004	A
20DAA	3.2097	0.8526	A

Different letters indicate significant difference among developmental stages using Tukey's LSD, $p<0.05$.

Appendix 15. Analysis of variance for relative abundance of *TaSUT2* transcript in palea during grain development.

Time	Estimate	Error	Group
10DAA	4.7667	0.8242	A
15DAA	4.0500	0.8242	A
30DAA	3.8000	0.8242	A
25DAA	3.7200	0.8242	A
20DAA	3.2850	1.0094	A
5DAA	2.4767	0.8242	A

Different letters indicate significant difference among developmental stages by Tukey's LSD, $p<0.05$.