Chapter 9

COTTON SEED DEVELOPMENT: OPPORTUNITIES TO ADD VALUE TO A BYPRODUCT OF FIBER PRODUCTION

Qing Liu, Danny J. Llewellyn, Surinder P. Singh and Allan G. Green
CSIRO Plant Industry, PO Box 1600, ACT 2601, Australia

INTRODUCTION

Unlike other major crops in which seeds provide most of the economic value, cotton seed is largely a by-product of more valuable fiber production and represents only approximately 15% of the farm gate value of the cotton crop. A cotton plant normally produces about 1.6 kg of seed for every kg of lint. Following ginning to remove the lint fiber, fuzzy cotton seed is either used directly as animal feed, or processed into four major products including oil, meal, hull and short linter fibers (Cherry and Leffler, 1984; O'Brien *et al.*, 2005). With about 21% oil and 23% protein by weight in the seed, however, this makes cotton the fifth largest oil crop in the world and the second most important potential source of plant proteins. As a by-product of fiber production, the continued large scale production of cotton seeds is assured because of the on-going demand for the fiber by the textile industry. Until recently, the quality of cotton seed has been relatively neglected as the main research priority has been devoted to enhancing fiber quality and fiber yield. This continual human selection for high yield and better quality of fiber has narrowed the genetic variation available for cotton seed quality improvement through breeding. An in-depth understanding of metabolic events that determine the overall components of the storage reserves in cotton seeds is therefore of vital importance for improving the yield, quality and ultimately the value of seed products and opens the possibility for significant value-adding by engineering novel attributes into the seed. While in this review we have focused primarily on the development of the cotton seed, other more studied species such as *Arabidopsis thaliana* can serve as a model for cotton and for completeness those other species are mentioned where data for cotton are limited.

EMBRYO AND SEED DEVELOPMENT ARE HIGHLY COORDINATED WITH FIBER PRODUCTION

As in other angiosperm species, seed development in cotton is triggered by the double-fertilization event that gives rise to the zygote and the triploid cell that gives rise to the endosperm tissue. At the same time that the zygote and endosperm are generated, the maternally derived seed coat begins to differentiate from the ovule outer integuments that surround the embryo sac. The seed coat plays a major role in protecting the embryo and transferring nutrients from the maternal plant to all parts of the seed. It also generates the seed fibers that provide a dispersal mecha-

nism for seeds in the wild. Individual epidermal cells within the expanding seed coat elongate rapidly at anthesis to produce the long lint fibers that thicken and desiccate into the fibers that are harvested from the seed at maturity. The development of the fiber is described in detail in Chapter 10. A schematic drawing of a typical cotton seed at maturity is shown in Figure 1.

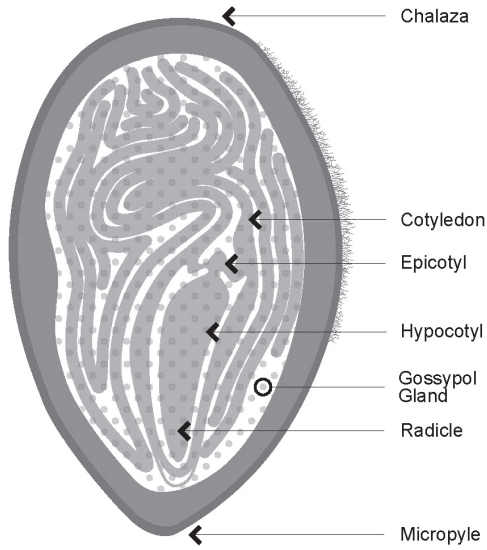

Figure 1. A schematic drawing of a typical cotton seed at maturity.

Embryogenesis

The development of the embryo can be generally classified into three overlapping stages, namely embryogenesis, maturation and desiccation (West and Harada, 1993; Turley and Chapman, 2010). Following fertilization, the endosperm proliferates to occupy most of the post-fertilization embryo sac and nourishes the embryo during early development (Schulz and Jensen, 1977). The zygote divides asymmetrically, giving rise to the apical cell that develops into the embryo proper and the suspensor structure (funiculus) that supports and feeds the embryo until it degenerates later in development (Goldberg *et al.,* 1994). The zygote undergoes its first cell division at 3-5 days post anthesis (DPA). The formation of the heart stage embryo occurs at 6-17 DPA. All embryonic organs and vascular tissues are differentiated and developed by 18 DPA at the torpedo stage. Embryos then continue to expand until approximately 25 DPA. The fiber and embryo share synchronous phases of development (Stewart, 1986). Rapid cell growth in both fiber and embryos occurs in the period of the first 25 DPA. Zygotic differentiation and growth through the globular, heart, and torpedo stages are well documented, but the chronology of the major events occurring during each phase is variable, likely due to different varieties and environmental conditions (Reeves and Beasley, 1935; Forman and Jensen, 1965; Dure, 1975). Cotton embryos have two large cotyledons that accumulate lipid, storage protein, and phytin, but no starch. In contrast to some other dicotyle-

donous plant species, such as legumes, cotton cotyledons remain thin, become highly convoluted, and wrap themselves around the embryo axis (Fig. 1).

The embryo maturation stage over the next 20 days (25-45 DPA) is characterized by an increase in the size and weight of the cotyledons and rapid accumulation of oil and storage proteins that are the major reserves of carbon and nitrogenous compounds needed for seed germination and early seedling growth, respectively (Reeves and Beasley, 1935; Forman and Jensen, 1965). During this embryo maturation period the fiber on the seed coat is depositing its thick secondary cell wall of almost pure cellulose (Stewart, 1986). The strengthening and lignification of the non-fiber seed coat epidermal cells also occurs at this time. Approximately 80% of the dry weight of the mature cotton seed kernel consists of storage lipids and proteins. Different from some other oilseed plants such as rapeseed (Murphy, 1993), the temporal accumulation of storage lipids in cotton seeds coincides with the accumulation of storage proteins (Chapman unpublished and referred to in review by Turley and Chapman, 2010).

At the final stage of cotton seed development after approximately 45 DPA, the funiculus begins to disintegrate and both seeds and fiber desiccate through lack of a vascular connection (Benedict *et al.*, 1976). The embryo reaches its maximum dry weight while its water content declines (Hughes and Galau, 1989; Turley and Chapman, 2010). Developmental arrest and the ability to withstand desiccation are the characteristics of cotton embryos at this phase and enable the embryo to remain in a quiescent state without undergoing precocious germination.

Endosperm Development

The endosperm nucleus divides freely within hours after fertilization forming a multinucleate mass surrounding the embryo. Cellular endosperm forms at about 10 DPA. The increase in its dry weight peaks about 24 DPA, then gradually decreases as the embryo expands and stabilizes by 35 DPA (Stewart, 1986). Little is known about the co-ordination between the developing embryo and surrounding endosperm and seed coat. There is emerging evidence in arabidopsis, at least, that the endosperm is not only a source of nutrients, but also acts as an integrator of seed growth and development through intercellular signal transduction pathways involving small secreted peptide ligands (Berger *et al.,* 2006; Fiume and Fletcher, 2012).

Molecular Regulation of Seed Development

Despite the obvious genetic control of cotton seed development and physiology, global gene expression programs involved in cotton embryogenesis, especially the early events following fertilization, are still largely unknown. In recent years arabidopsis and legumes have been used as model plants for genome-wide expression profiling and have provided some insights into the processes that may also occur in cotton (McElver *et al.,* 2001; Tzafrir *et al.,* 2004).

The developmental events that culminate in the production of a mature seed from a single-cell zygote are precisely coordinated and relatively conserved in the majority of angiosperm species (Goldberg *et al.,* 1994). The recent characterization of regulatory genes identified through chemical and insertional mutagenesis has provided the first glimpses of the developmental pathways involved in zygotic embryogenesis. By searching the Seed Genes database (http://www.

seedgenes.org) and recent literature, Hsu *et al.* (2010) found that 339 non-redundant genes are required for proper embryo formation. Of these, 108 likely encode plastid-targeted proteins, and 19 are necessary for development of globular stage embryos.

A large number of genes with complex expression patterns are also involved in somatic embryogenesis (Zeng *et al.,* 2007). Various signal transduction pathways are involved in activating/repressing numerous gene sets, many of which are yet to be identified and characterized. Molecular and physiological events leading to seed formation are therefore still far from being completely understood. A large number of genes required for seed development were revealed via T-DNA insertional mutagenesis in arabidopsis (McElver *et al.,* 2001) and surprisingly, not all were embryo-specific in their pattern of expression. This is consistent with a requirement for their basal functions throughout the life cycle of the plant (Tzafrir *et al.,* 2004).

While little has yet been done in cotton to understand the gene networks and signaling pathways necessary for embryogenesis it is unlikely that they will dramatically depart from that seen in model plants. Early embryogenesis in arabidopsis is regulated by transcription factors, signal transduction pathways mediated by kinases, and proteins that establish and maintain auxin hormone gradients (Willemsen and Scheres, 2004). For example, WUSCHEL-related homeobox (WOX) transcription factor genes mark cell fate decisions during early embryogenesis (Jenik and Barton, 2005). This gene is necessary for cell divisions that form the apical embryo domain (Haecker *et al.,* 2004). PIN-formed (PIN) genes encode transporter-like membrane proteins that are important for regulating auxin transport and mutations in *PIN1* and *PIN7* disrupt the establishment of the embryogenic apical-basal axis (Steinmann *et al.,* 1999). Seed development is also controlled epigenetically and DNA methylation, for example, is critical for plant embryogenesis and for seed viability. Arabidopsis plants with loss-of-function mutations in *MET1* (METHYLTRANSFERASE1) and *CMT3* (CHROMOMETHYLASE3) produce embryos that develop improperly and have reduced viability (Xiao *et al.,* 2006). Genes that specify embryo cell identity are mis-expressed and auxin hormone gradients are not properly formed in abnormal *met1* embryos. A fundamental aspect of the transition phase of embryogenesis, when apical and basal organ fates are established, is the repression of the root-promoting gene cascade in the apical embryo domain. This is provided by the activity of the TOPLESS and TOPLESS-RELATED proteins, co-repressors that function together with AUX/IAA proteins in auxin-mediated transcriptional repression (Szemenyei *et al.,* 2008).

Several transcription factors such as LEAFY COTYLEDON1 (LEC1), LEAFY COTYLEDON2 (LEC2), FUSCA3 (FUS3) and abscisic acid insensitive3 (ABI3) have been identified as master regulators of seed development and maturation, activating genes encoding seed proteins that define each phase of embryo development (Abid *et al.,* 2010). LEC proteins stimulate abscisic acid (ABA) levels and activate genes that repress gibberellic acids (GA) levels, contributing to the high ABA to GA ratio characteristic of the maturation phase. Ectopic expression of all three *LEC* genes, *LEC1*, *LEC2* and *FUS3*, caused cells in vegetative and reproductive tissues to adopt characteristics of maturation phase embryos (Lotan *et al.,* 1998; Santos-Mendoza *et al.,* 2005; Braybrook and Harada, 2008). Ectopic expression of *LEC1* has demonstrated that this gene is required for the maintenance of the suspensor cell fate, the specification of cotyledon identity, the initiation and maintenance of the maturation phase, the promotion of embryogenic

cell identity and division and the suppression of germination (Casson and Lindsey, 2006). Seed storage protein gene expression is controlled by LEC1 through the regulation of *ABI3* and *FUS3* expression (Kagaya *et al.,* 2005). Yamagishi *et al.* (2005) isolated *TANMEI* (*TAN*) that controls various aspects of both early and late phases of embryo development. Analysis of *tan* mutants indicates that this gene plays a role in both the morphogenesis and maturation phases of embryogenesis. Moreover, *tan* mutants share many characteristics with *lec* mutants suggesting that *TAN* has overlapping functions with the *LEC* genes. For example, both *tan-1* and *lec* mutations cause desiccation intolerance and defects in storage protein and lipid accumulation.

MicroRNAs (miRNAs) are 21-nucleotide single stranded RNA molecules that act by binding complementary target mRNAs to promote their cleavage or interfere with their translation (for review, see Voinnet, 2009). During early embryogenesis of arabidopsis, specific miRNAs down-regulate one or more miRNA targets to both promote the repression of maturation and, in the advanced embryogenic and developmental stages, a reduction in those miRNAs leads to the induction of maturation (Willmann *et al.,* 2011). DICER-LIKE1 (DCL1) that is required for miRNA biogenesis was found to be critical for multiple embryonic cell differentiation events as early as the eight cell stage. Genome wide transcript profiling revealed that DCL1 was required for the early embryonic repression of nearly 50 miRNA targets among which two redundant SQUAMOSA PROMOTER-BINDING PROTEIN-LIKE (SPL) transcription factors, SPL10/SPL11, regulated by miR156, were most depressed in eight-cell *dcl1* embryos (Nodine and Bartel, 2010). Molecular analyses of miRNAs in cotton (Zhang *et al.*, 2007; Barozai *et al.,* 2008; Wang *et al.*, 2012) indicate that most miRNA families and their specific gene targets are highly conserved between cotton and arabidopsis so will likely carry out the same roles in seed development.

Somatic Embryogenesis as a Model for Embryo Development

In a parallel with the developmental pathway of zygotic embryos in seeds, somatic embryos can be induced in cotton *in vitro* from a group of meristematic cells that go through globular, heart shaped, torpedo shaped and cotyledonary stages (Price and Smith, 1979; Shoemaker *et al.,* 1986). The zygote is intrinsically embryogenic, while somatic embryogenesis requires the induction of embryogenic competence. Despite these differences, some common features between zygotic and somatic embryogenesis have been reported (Trolinder *et al.,* 1987, 1988, 1989). One of the major obstacles to understanding in detail the events that govern early embryo formation in cotton is the relative inaccessibility of the young embryo to experimental manipulation, particularly at the early stages of embryogenesis involving only one or a few cells. Somatic embryogenesis offers an alternative approach that circumvents this problem in some respects as large quantities of early embryos can be simultaneously induced and studied at the molecular or biochemical level.

Gene expression associated with somatic embryogenesis in somatic embryogenic responsive cotton cultivars including Coker 201 and CCR124 was analyzed using a suppression subtractive hybridization (SSH) approach (Zeng *et al.*, 2006; Wu *et al.,* 2009). Genes encoding transporters such as lipid transfer protein (LTP) were found to be the major group of genes responsive to

somatic embryogenesis. Expression of LTP genes is strongly associated with the first differentiated tissue of somatic embryos and exerts a regulatory role in controlling cell expansion during embryo development (Dodeman *et al.,* 1997). This induction of different transporters was followed by induction of genes encoding proteins involved in regulating transcription and post-transcriptional processing, including Zinc Finger-HD homeobox genes, ethylene-responsive transcriptional co-activators, a putative RNA helicase, a SCARECROW (SCR)-like gene, and genes encoding nonsense-mediated mRNA decay proteins and high mobility group proteins. Homeodomain genes play a key regulatory role in pattern formation and differentiation in all multi-cellular organisms. SCR is critical in regulating the asymmetric division of the cortex/endodermis initial and radial patterning of tissues in both the root and shoot during embryogenesis (Di Laurenzio *et al.,* 1996; Wysocka-Diller *et al.,* 2000).

METABOLISM IN THE DEVELOPING SEED AND ITS MANIPULATION THROUGH BIOTECHNOLOGY

A hallmark of plant embryos is the accumulation of storage reserves and secondary metabolites. Cotton seed is a rich source of oil and proteins that have been the subject of many investigations and evaluated as a wholesome, nutritious and versatile ingredient in animal feed as well as human food products (review by Cherry, 1983; Alford *et al.,* 1996; Liu *et al.,* 2009). Cotton seed is also characterized by the accumulation of lysigenous gossypol glands containing terpenoid aldehydes, principally gossypol that is toxic to non-ruminant animals and humans (Fryxell, 1968; Stipanovic *et al.,* 1977). Recent advances in the understanding of the biochemical, cellular and molecular mechanisms underlying biogenesis of seed storage compounds and secondary metabolites, coupled with the cloning of many of the genes involved in these processes, have facilitated the production of designer cotton seeds with improved nutritional benefits and enhanced functional properties that will add value to this by-product of fiber production.

Carbohydrate Metabolism

Developing cotton fruits (bolls) are very strong carbohydrate sinks. The fruit can increase in weight by as much as 15% per day (Schubert *et al.,* 1986). During the early stage of seed development, sucrose derived from the photosynthetic leaves is delivered to the seed coat and endosperm, via the funiculus, and it is either cleaved by sucrose synthase (Sus) producing UDP-glucose and fructose or hydrolysed by invertase into glucose and fructose (Hendrix, 1990; Ruan *et al.,* 2008). The comparison of relative activities of sucrose synthase and invertase in the seed coats at this stage suggested that the Sus seems to predominate over invertase (Hendrix, 1990; Ruan *et al*., 2008). UDP-glucose could be transported to fiber cells and directly used in cellulose biosynthesis (Benedict *et al.,* 1976; Haigler *et al.,* 2001). However, much of the hexoses are used to synthesize starch in the seed coat and endosperm, at least transiently (Hendrix, 1990; Ruan *et al*., 2008).

Cotton embryos accumulate starch beginning at 20 DPA when starch in the seed coat begins to decline. At about 40 DPA, the starch in the embryo starts to disappear and approximately the

same quantity of complex sugars appears as galactosides including raffinose and stachyose. By about 45 DPA, carbon importation into the seed has ceased and re-arrangement of non-structural carbohydrates becomes a major metabolic activity driving seed maturation (Benedict *et al.*, 1976). Raffinose is synthesized by the transfer of galactose to sucrose from galactinol, while a second galactose addition produces stachyose. Mature cotton seeds contain up to 10% by weight of these soluble storage carbohydrates (Muller and Jacks, 1983; Mellon and Cotty, 1999). Raffinose and stachyose disappear rapidly early after germination and they serve as readily mobilizable carbon sources before the lipid utilizing enzymes become fully active (Bortman *et al.*, 1981; Doman *et al.*, 1982; Miernyk and Trelease, 1981; Trelease *et al.*, 1986).

It has been well established that Sus plays an important role in young cotton ovule carbohydrate partitioning and that sucrose phosphate synthase and the galactoside synthesizing enzymes are important in carbohydrate partitioning in maturing cotton seeds (Hendrix, 1990; Ruan, 2005). In early seed development, Sus was localized in the cellularising endosperm, but not in the heart shaped embryo at 10 DPA (Ruan *et al.*, 2008). In developing cotton embryos, Sus activity was found to be about five-fold that of vacuolar invertase and more than 10-fold that of alkaline invertase (Ruan *et al.*, 2003), indicating a key role for Sus in degrading incoming sucrose during early cotton seed development. Suppressing the expression of Sus in the cotton seed coat led to a fiberless phenotype (Ruan *et al.*, 2003), whereas silencing its expression in the filial tissue resulted in stunted and unviable seeds and loss of transfer cells (Ruan *et al.*, 2003; 2008). On the other hand, elevation of Sus activity by over-expressing a potato *Sus* gene in transgenic cotton resulted in an increased seed number (Xu *et al.*, 2011). The higher Sus activity in the filial tissues might significantly enhance seed sink strength (Pugh *et al.*, 2010) and consequently reduce seed abortion to increase the number of seeds set.

Oil Biosynthesis and Accumulation in Cotton Seeds

Oil is one of the major compounds synthesized and deposited during cotton seed development. It is composed almost entirely of triacylglycerol (TAG) molecules that consist of three fatty acids (FAs) bound to a glycerol backbone. FA biosynthesis and TAG assembly are highly compartmentalized processes that are located in several different sub-cellular organelles. *De novo* biosynthesis of FAs occurs in the stroma of plastids in the developing seeds. FAs are then exported to the cytoplasm in the form of acyl-CoA thioesters, and some of these FAs may also be modified while attached to phospholipids on the surface of endoplasmic reticulum (ER) prior to TAG assembly and storage in the oil bodies or oleosomes (see review by Browse and Ohlrogge, 1995).

In most plant species, including cotton, the sugars providing the carbon skeleton for lipids synthesized in the seed are first transported into the endosperm from the source tissues via the seed coat and are then absorbed by the embryo. The sugars include sucrose, glucose and fructose. The cleavage products of sucrose generated by Sus and invertase are utilized through both the cytosolic and the plastidic glycolytic pathways. Multiple transporters, including the glucose-6-phosphate (Glc-6-P) translocator, the triose phosphate translocator, and the phosphoenolpyruvate (PEP) translocator, which are localized in the plastid membrane, exchange intermediates

that are generated during glycolysis from the cytosol to the plastid (Fischer and Weber, 2002). The PEP and pyruvate generated are utilized in plastids as the carbon source for FA biosynthesis. Malonyl-CoA, produced from pyruvate, serves as a two-carbon unit donor in incremental FA synthesis catalysed by a series of fatty acid synthases (Browse and Ohlrogge, 1995).

While little detailed molecular analyses have been done in cotton, Bourgis *et al.* (2011) recently compared the transcriptome and metabolomes of oil accumulating developing mesocarp of oil palm (*Elaeis guineensis*) and sugar accumulating developing mesocarp of date palm (*Phoenix dactylifera*) to gain an understanding of the reasons for the large difference in carbon partitioning between these two closely related palm species. Among all the genes involved in fatty acid biosynthesis and TAG assembly, only transcripts encoding the 18 plastidial fatty acid biosynthetic enzymes that are involved in the conversion of pyruvate to fatty acids were found to be significantly up-regulated in oil palm by, on average, 13 times. Plastidial transporters for hexose, pentose, triose phosphate, and phosphoenolpyruvate were all strongly up-regulated in oil palm, pointing to increased capacity for carbon flow into the plastid. Surprisingly, most enzymes of TAG assembly and cytosolic glycolysis were not very different between the two species, indicating that it is FA biosynthesis and transfer rather than TAG assembly that exert the most control over oil accumulation in palm and possibly other plants.

Sucrose and glucose are the major source of carbon provided by maternal tissues to developing seeds. In embryos and endosperm of four oilseeds, including *B. napus*, *Ricinus communis*, *Euonymus alatus* and *Tropaeolum majus*, Sus ESTs were 20-40 fold higher than neutral invertases (Tronocoso-Ponce *et al.,* 2011). While there are many different levels of regulation of metabolic flux through pathways other than at the transcriptional level, these EST data suggest Sus may play a key role in generating hexoses during oil accumulation in embryonic cotyledons. Such a result appears to be consistent with the biochemical analysis in developing cotton seeds (Ruan *et al.*, 2003; 2008).

In the plastids of all plants, acetyl-CoA carboxylase (ACCase) catalyses the first committed step in the pathway of oil synthesis by addition of a carboxyl group to acetyl-CoA to form malonyl-CoA. The malonyl group is then transferred from CoA to an acyl-carrier protein (ACP), which serves as the carrier for the growing FA chain. The acetyl-CoA condensing enzyme, ketoacyl-ACP synthase III (KASIII) catalyses the elongation step of acetyl-CoA and malonyl-ACP to yield acetoacetyl-ACP, the first step in the fatty acid elongation pathway (Clough *et al.,* 1992). The repeated process of adding two-carbon units onto the elongated fatty acid chain is catalysed by KASI leading to the formation of palmitoyl-ACP (C16:0-ACP). KASII catalyses the elongation of palmitoyl-ACP to stearoyl-ACP (C18:0-ACP). A soluble stearoyl-ACP Δ9-desaturase (SAD) introduces the first double bond into stearoyl-ACP to convert it to oleoyl-ACP in the plastid. Cotton seed oil accumulates only about 2% stearic acid as much of stearate is converted to oleate *via* SAD. SAD enzymes are encoded by a multigene family consisting of at least five genes per diploid genome in cotton (Liu *et al.,* 1996; 2009). The growing saturated fatty acyl chain and the monounsaturated oleate are cleaved off the ACP by a substrate specific thioesterase enzyme, FatB or FatA, enabling them to exit the plastid into the cytoplasm. On ER membranes, oleic acid becomes associated with Phosphatidylcholine (PC) and can be further modified by a membrane-bound omega-6 (or Δ12) fatty acid desaturase, FAD2. FAD2 intro-

duces a double bond into oleic acid to form the linoleic acid that accounts for more than 50% of total fatty acids in cotton seed oil. There are at least four different genes encoding cotton FAD2, with *ghFAD2-1* playing a major role in the production of linoleic acid in cotton seed oil (Liu *et al.,* 1999a,b; Pirtle *et al.,* 2001; Zhang *et al.,* 2010). The expression of *ghFAD2-1* is seed-specific and reaches its highest level at the middle maturity stage, between 25-35 DPA, before drastically declining when seeds approach maturity (Liu *et al.,* 1999a). In cotton, oleic acid is also the precursor for a group of carbocyclic fatty acids, i.e. dihydrosterculic acid, sterculic acid and malvalic acid, which account for less than 1% of total FAs. In a developing cotton seed, these carbocyclic fatty acids are exclusively synthesized in the embryo axis, but not in cotyledons (Wood, 1986). A fatty acid methyltransferase, known as cyclopropane FA synthase (CPA-FAS), was identified as the first enzyme converting oleic acid to dihydrosterculic acid using *S*-adenosylmethionine as the methyl donor (Bao *et al.,* 2003; Liu *et al.*, 2004; Yu *et al.,* 2011).

The FAs formed may then be incorporated into membrane and storage lipids *via* the Kennedy pathway by the sequential esterification of glycerol-3-phosphate through the action of glycerol-3-phosphate acyltransferase (GPAT) to form lysophosphatidic acid, followed by 1-acyl-sn-glycerol-3-phosphate acyltransferase (LPAAT) to form phosphatidic acid (PA). Dephosphorylation of PA by phosphatidic acid phosphatase results in the formation of diacylglycerol (DAG), which is then acylated to form TAG by a DAG acyltransferase (DGAT) (Browse and Ohlrogge, 1995). The recently discovered enzyme, phosphatidylcholine:diacylglycerol cholinephosphotransferase (PDCT) is required for the efficient biosynthesis of polyunsaturated fatty acid during TAG accumulation in seeds (Lu *et al.,* 2009). PDCT catalyses the inter-conversion of PC and DAG by transferring the phosphocholine headgroup between these two molecules, thus enabling more oleic acid be desaturated by the FAD2 enzyme. TAG can also be formed in plants *via* an acyl-CoA independent pathway, catalyzed by phospholipid:diacylglycerol acyltransferase (PDAT) (Dahlqvist *et al.,* 2000). Finally TAGs are stored in seeds in specialized oil body structures. Chapman studied the enzyme activities of some of these TAG assembly enzymes in developing cotton seed and found LPAAT and DGAT are highest in cotton embryos between 25 to 35 DPA (Chapman unpublished data, referred by Turley and Chapman, 2010).

Metabolite analysis (Perry *et al.,* 1999) and other studies (Zheng *et al.,* 2008) suggest that DGAT may be one of the rate-limiting steps in plant seed lipid accumulation. Over-expression of the arabidopsis *DGAT1* in wild type plants led to an increase in seed oil content and seed weight (Jako *et al.,* 2001). Subsequently, *DGAT* expression has been genetically manipulated to produce crops with increased oil content (Lardizabal *et al.,* 2008; Weslake *et al.,* 2008; Zheng *et al.,* 2008; Taylor *et al.,* 2009). The major reduction in oil content in seeds of arabidopsis with RNAi-mediated silencing of *DGAT1* and *PDAT1* indicates that these enzymes have overlapping functions in seed TAG synthesis, and also led to the discovery that TAG synthesis or other functions of DGAT1 and PDAT1 are essential for normal embryo and pollen development (Zhang *et al.,* 2009). TAG synthesis clearly plays an important role not only as a storage reserve, but also in normal growth and development.

In arabidopsis LEC1 and WRI1 have been identified as two key transcription factors involved in the regulation of oil accumulation (Baumlein *et al.,* 1994; Cernac and Benning, 2004) and over-expression of either or both of *LEC1* and *WRI1* resulted in oil increases without negative

impacts on grain yield (Shen *et al.,* 2010). In the developing seeds, carbon derived from photosynthesis is partitioned into different storage compounds. The carbon flux into oil may be in competition with other metabolic pathways, such as starch. It has been recently demonstrated that redirecting carbon partitioning by genetic down-regulation of *APS1*, a major catalytic isoform of the small subunit of ADP-glucose pyrophosphorylase involved in starch biosynthesis, in combination with up-regulation of oil biosynthesis through over-expressing *WRI1*, could substantially increase TAG accumulation and therefore lead to increased energy density of the seed biomass (Sanjaya *et al.,* 2011). However, this may not be applicable to cotton as it does not accumulate much starch in mature seeds.

Nutritional Enhancement of Cotton Seed Oil

Other than the fiber, oil is the most valuable product of a cotton seed, and is widely used as a cooking oil and an ingredient in marinades, dressings, pastries, margarines, and shortenings. Furthermore, oil is by far the most efficient form for energy storage, since it contains more than twice the energy on a dry weight basis than can be stored in starch or proteins. Consequently whole cotton seed has been regarded by the dairy industry as a special feed ingredient with advantageous energy and dietary fiber properties required by the high-producing dairy cow. There is increasing emphasis on greater energy density in cotton seed through increasing seed oil content at the expense of carbohydrate concentrations (Coppock *et al.,* 1987; O'Brien *et al.,* 2005).

In cotton seed, the ratio of oleic and linoleic acids is largely determined by the seed-specific enzyme, ghFAD2-1, which converts about 80% of oleic acid to linoleic acid in developing cotton seed. High levels of linoleic acid have made cotton seed oil prone to auto-oxidation. Partial hydrogenation is often used to extend the oil's stability and shelf life, producing various levels of *trans* fatty acids that are recognized to have cholesterol-raising properties and associated increased risk of cardiovascular heart disease (Mozaffarian *et al.*, 2006). Seed-specific down-regulation of *ghFAD2-1* using an RNAi approach was able to raise oleic acid levels up to 78%, at the expense of linoleic acid that was reduced to 4% (Liu *et al.,* 2002). Cotton seed oil also contains the highest level of palmitic acid among commodity vegetable oils produced in the temperate climates. Palmitic acid is widely reported to raise total plasma cholesterol and low-density lipoprotein cholesterol levels (Kris-Etherton *et al.,* 1993). The small amount of carbocyclic fatty acids that exist in cotton seed oil are potent inhibitors of mammalian fatty acid desaturases and are therefore believed to be nutritionally undesirable (Shenstone and Vickery, 1961; Cherry, 1983). Significant reduction of these nutritionally undesirable fatty acids would therefore increase the health appeal of cotton seed oil and competitiveness relative to other vegetable oils. In the subsequent experiments of Liu and co-workers (Liu *et al.*, 2008b), two other genes, *ghFatB* and *ghCPA-FAS*, were simultaneously down-regulated together with *ghFAD2-1*. The transgenic cotton plants had reduced palmitic acid and carbocyclic fatty acids and a further increase of oleic acid, representing potentially enhanced nutritional value.

Seed-specific RNAi mediated gene silencing of *ghSAD-1* led to a raised stearic acid level up to 40% on the single seed basis, accompanied by a reduction of all other fatty acids in cotton seed oil: i.e., linoleic acid was reduced to 38% from the normal 56%; palmitic acid to 17% from 26%; and oleic acid to 10% from 15% (Liu *et al.,* 2002). Such high-stearic oil was found to have a markedly increased melting point compared to the conventional cotton seed oil control

(Qing Liu, unpublished data). Different from palmitic acid, its shorter chain saturate counterpart, stearic acid does not raise cholesterol and therefore does not impact heart health negatively (Mensink and Katan, 1992). Because of its high oxidative stability and high melting point, cotton seed oil with raised stearic acid level could be used to manufacture *trans*-free margarine and as a substitute for cocoa butter in confectionery.

Albeit lacking omega-3 fatty acids, traditional cotton seed oil does contain high levels of linoleic acid that could provide a significant endogenous substrate pool for the Δ6 desaturase that produces γ-linolenic acid (C18:$3^{\Delta6,9,12}$, GLA) and the Δ15 desaturase that produces α-linolenic acid (C18:$3^{\Delta9,12,15}$, ALA) that are found in other plants. The Δ6 desaturase could also catalyse further desaturation of ALA to produce stearidonic acid (C18:$4^{\Delta6,9,12,15}$, SDA), a precursor for the production of omega-3 long chain polyunsaturated fatty acids (LC-PUFA). Transgenic expression of a Δ6 desaturase from *Echium plantagineum* and a Δ15 desaturase from *Brassica napus* has led to the accumulation of 25% GLA and 35% ALA in cotton seed oil, respectively (Liu *et al.*, 2008a). Up to 6% SDA, in addition to both ALA and GLA, were observed in homozygous F_1 seeds following the hybridization of these two transgenic genotypes. Such genetically modified cotton seed oil may provide an alternative source for these C18 polyunsaturated fatty acids that are currently obtained from herbal oils such as those from evening primrose, borage and black current. The transgenic cotton seeds expressing the entire omega-3 LC-PUFA pathway may also be able to replace the current marine sources for omega-3 LC-PUFA, such as eicosapentaenoic acid (20:$5^{\Delta5,8,11,14,17}$, EPA) and docosahexaenoic acid (22:$6^{\Delta4,7,10,13,16,19}$, DHA).

Improving cotton seed oil quality may not be limited to just increasing its nutritional value. The accumulation of vernolic acid, an epoxygenated FA with valuable industrial applications, increased to up to 17% of total fatty acids in the seed oil of transgenic cotton plants expressing a Δ12 fatty acid epoxygenase derived from *Crepis palaestina* (Zhou *et al.*, 2006). However this level of vernolic acid in cotton seed oil is still considerably less than the 70% found in *C. palaestin*a, the source of the epoxygenase transgene. It is apparent that metabolic bottlenecks are limiting the full potential for cotton seed in accumulating this unusual fatty acid at an economically viable level, and further research would be required to overcome this bottleneck. The production of unusual fatty acids in oilseed crops in general has been a great challenge because of the apparent complexity of the biosynthesis pathways for these fatty acids (Napier, 2007).

Synthesis and Accumulation of Seed Storage Proteins in Cotton Seeds

In higher plants, seed storage proteins are synthesized on the rough ER, using amino acids directly taken up by the embryo, or obtained after transamination reactions. Subsequently, they are transported into protein storage vacuoles by a vesicle-mediated pathway (Jolliffe *et al.*, 2005). In cotton seed, two major classes of storage proteins are globulins and albumins, which differ in their solubility properties. Both globulins and albumins are synthesized and compartmentalized in storage protein vacuoles during cotton seed maturation (Dure and Chlan, 1981). Globulins can be further classified based on the sedimentation rate of their aggregated forms into the 7S vicilins (or α-globulin) and 11/12S legumins (or β-globulin; Youle and Huang, 1981). In a recent

survey of the most abundant cotton seed storage proteins, nearly all of the proteins identified belong to the vicilin and legumin families, comprising 60-70% of the total seed proteins (Hu *et al.,* 2011).

Vicilin A and Vicilin B, which share 72% amino acid similarity, represent the first discovered cotton seed storage proteins (Chlan *et al.,* 1987). Each vicilin is encoded by a single-copy genes in the diploid *Gossypium* genomes, and their corresponding homeologous copies were identified in allotetraploid cotton. Similarly there are two legumin isomers, Legumin A and Legumin B, which are more diverged compared to the vicilin gene family, sharing only 58.5% amino acid similarity. As with the vicilins, each legumin is encoded by a single-copy gene in the diploid cotton genomes.

The expression of seed storage proteins genes is regulated temporally and spatially in cotton seeds (Galau *et al.,* 1983; Chlan *et al.,* 1987; Galau *et al.,* 1992). Despite the existence of mainly single genes encoding each protein in the diploid genome, the mRNA pool in developing cotton seeds was composed of 30% legumin mRNAs, 15% vicilin mRNAs, and 2% 2S albumin mRNAs (Hughes and Galau, 1989; Galau *et al.,* 1992). Therefore, regulatory sequences from genes encoding seed storage proteins represent a valuable source of promoters that could be utilized to drive the expression of transgenes in a seed-specific manner. The promoter region of α-globulin gene B was isolated and identified as highly seed-specific (Sunilkumar *et al.,* 2002), and it has been successfully used for seed-specific reduction of toxic gossypol levels in cotton seed (Sunilkumar *et al.,* 2006).

Cotton seed protein contains large amounts of arginine, especially compared to some legume species such as soybean (Capdevila and Dure, 1977). Arginine has been shown to slow down cancer progression (Lowell *et al.,* 1990), to act as a principal regulator of blood pressure, and to cause a relaxation of cardiovascular smooth muscle cells following conversion to nitric oxide (Moncada and Higgs, 1993). However, cotton seed is severely deficient in lysine and slightly deficient in isoleucine and the sulphur amino acids, including methionine and cysteine, compared to other major oilseeds, such as soybean and canola (Capdevila and Dure, 1977). Additionally, when pigment glands in the cotton seed are disrupted during processing, the free gossypol can bind to the epsilon amino group of lysine and reduce its availability to below acceptable levels for animal nutrition (Zarins and Cherry, 1981; Calhoun *et al.,* 1995).

Genetic improvement of cotton seed storage protein and amino acid profiles are clearly necessary if it is to be used as a source of protein for non-ruminant animals or humans, but little if any progress has been documented in the literature. The feasibility of raising lysine content in a plant seed has been demonstrated in maize in recent years (Huang *et al.,* 2006; Frizzi *et al.,* 2008), and this could also be applicable to cotton. To imitate the natural high lysine maize *opaque* mutant with its reduced level of the seed storage prolamine α-zein, RNAi was used to specifically suppress α-zein production in transgenic maize kernels, resulting in a doubling of the lysine content from 2400 to 4800 ppm (Huang *et al.,* 2006). To further enhance the accumulation of free lysine, the bifunctional enzyme lysine-ketoglutarate reductase/saccharopine dehydrogenase (LKR/SDH), which is responsible for lysine catabolism, was also expressed in transgenic maize and increased lysine content about 30 fold. An additional bottleneck for lysine accumulation was identified as the lysine feedback effect, which was by-passed by express-

ing a lysine feedback-insensitive enzyme from *Corynebacterium glutamicum*, CordapA. This resulted in further enhancement of lysine levels up to 100 fold in maize endosperm relative to non-transgenic controls (Frizzi *et al.*, 2008).

Cotton seed is also deficient in sulphur-containing amino acids because the sulphur-rich proteins, such as albumin, constitute a low fraction of the total seed proteins (Galau *et al.*, 1992; Hu *et al.*, 2011). The sulphur amino acids, such as methionine, in cotton seeds could be genetically enhanced by either over-expressing a sulphur-rich protein or reducing the relative abundance of endogenous sulphur-poor proteins, such as globulins. It has been demonstrated in rapeseed and lupin seeds that methionine content can be increased substantially by the introduction of foreign genes encoding naturally sulphur-rich proteins, such as the 2S albumin from Brazil nut that contains 18% methionine and 8% cysteine (Altenbach *et al.*, 1992; Molvig *et al.*, 1997). These approaches take advantage of the plant's homeostasis mechanisms such that seeds generally appear to compensate for a shortfall of a major protein by accumulating other seed proteins to maintain a relatively constant protein content. For example, the genetic reduction of β-conglycinin in soybean resulted in a concomitant increase in the accumulation of glycinin, which contains higher levels of sulphur amino acids (Kinney *et al.*, 2001). However, despite the observed increase of methionine in transgenic seeds, total seed sulphur amino acids remained virtually unchanged relative to the control plants. Molvig *et al.* (1997) expressed the sunflower seed albumin (SSA) protein that is rich in methionine (16%) and cysteine (8%) in transgenic lupin, resulting in a 94% increase in methionine accompanied by a 12% reduction of cysteine, another sulphur-rich amino acid. The unexpected decrease in cysteine suggested that synthesis of the transgenic protein simply caused a redistribution of limited sulphur resources. A substantial increase in the bioavailable sulphur pool would be necessary for any real gain in these nutritionally desirable sulphur-rich amino acids.

The sulphur-rich 2S albumins, unfortunately, have been recognized as potent allergens and correlated with potentially lethal anaphylaxia in numerous plant seeds such as Brazil nuts, peanuts and sunflower seeds (see review by Herman and Burks, 2011). Despite some early efforts in biochemical characterization of cotton 2S albumin, much remains poorly understood about the potential allergenic effects of cotton seed storage proteins (Youle and Huang, 1979). Such an issue needs to be addressed before genetic enhancement of 2S albumin in cotton seeds could be undertaken.

A class of 18 proteins termed as late embryogenesis abundant (LEA) proteins was found to be highly up regulated in the seed desiccation period in cotton (Hughes and Galau, 1989). It is hypothesized that some of these LEA proteins are functionally involved in eliciting desiccation tolerance in the seed, and their synthesis is correlated with abscission of the funiculus that terminates nutrition and water transport to the seed from the mother plant (Hughes and Galau, 1989; Turley and Chapman, 2010).

Oil Body Proteins: Oleosins

Oleosins are a class of small proteins associated with the oil body membrane in plant seeds (Huang, 1992; Hughes *et al.*, 1993; Chapman *et al.*, 2012). They play dual physiological roles,

acting as protectors for stabilizing the oil bodies in developing and mature seeds and as the recognition signal for lipase binding in germinating seeds. Oleosins are alkaline and hydrophobic proteins having three domains including amphipathic N and C termini and a central hydrophobic domain that is highly conserved and could penetrate through the phospholipid monolayer into the oil body matrix. The N and C termini of the oleosin polypeptide are located on the oil body surface and interact with the phospholipid membrane, forming a highly stable structure that surrounds the oil body in an amphipathic shell (Huang, 1992). Two distinct genes encoding oleosin, *MatP6* and *MatP7* that are 77% identical, were expressed during the maturation and post-abscission stages of cotton embryogenesis (Hughes and Galau, 1989; Hughes *et al.,* 1993).

Taking advantage of the oleosin's capability for anchoring onto the surface of oil bodies, artificial oil bodies (AOBs) have been successfully constituted and used for purification, refolding, and immobilization of recombinant proteins in transgenic oilseed plants. The heterologous protein fused to the N/C-terminus does not alter the functional domains of the oleosin, and the fusions can still anchor normally onto the surface of oil bodies. Since oil bodies are of low density and free of contaminating proteins, they can be separated from other cellular components simply by flotation centrifugation. The desired protein along with the oil body fraction can be readily removed without the need for costly and time consuming chromatographic steps for protein purification. The desired protein can then be released from oil bodies by using a site-specific protease to cleave at a sequence engineered into the recombinant protein. This property has been exploited for numerous "molecular farming" applications, as reviewed recently by Boothe *et al.* (2010). For example, fusion proteins of oleosin with either a human precursor insulin (Des-B30) (Nykiforuk *et al.,* 2005) or a human insulin-like growth factor 1 (hIGF-1) have been successfully expressed in arabidopsis seeds (Li *et al.,* 2011). Similar strategies could be used to generate valuable recombinant proteins in cotton.

Gossypol Synthesis and Accumulation in Cotton Seeds

Lysigenous glands in cotton plants, including cotton seed tissues, contain terpenoid aldehydes. These are sesquiterpenes (C15) derived from a cytosolic branch of terpenoid metabolism via the mevalonate pathway (Stipanovic *et al.,* 1999; Cai *et al.,* 2010), and they provide a defense against herbivory. Farnesyl diphosphate (FPP) is generated as the linear carbon skeleton of the sesquiterpenes, and its cyclisation (catalysed by a terpene cyclase enzyme, (+)-δ-cadinene synthase, to form (+)-δ-cadinene) is the first committed step in gossypol biosynthesis (Chen *et al.,* 1995). (+)-δ-cadinene is then hydroxylated at the C-8 position (leading to 8-hydroxy-(+)-δ-cadinene) through the action of a cytochrome P450 enzyme, (+)-δ-cadinene-8-hydroxylase (CYP706B1). Subsequently, 8-hydroxy-(+)-δ-cadinene is converted to desoxyhemigossypol (dHG) by a yet uncharacterized process and further oxidized by one electron into hemigossypol. Finally, gossypol is formed by a phenolic oxidative coupling of two molecules of hemigossypol, catalysed by a hydrogen peroxide-dependent peroxidase enzyme (Dewick, 2009).

Gossypol is the predominant sesquiterpenoid formed in cotton seed, with only traces existing of dHG and hemigossypol (Cai *et al.,* 2010). Gossypol occurs in either free or bound to proteins, and the former is toxic. Gossypol in its unbound form causes anorexia, slow growth, and in-

creased fat deposition in liver tissue when fed to fish in excess (Wood and Yasutake, 1956). Gossypol is particularly toxic to non-ruminant animals and has inhibitory effects on male fertility when cotton seed is used for feeding either directly or as a meal following oil extraction. During the oil extraction process, gossypol is deactivated in the meal through moist heating, which causes the formation of a double bond between the ε-amino group of lysine and the aldehyde group. Although effective in reducing the toxicity of gossypol, the binding of gossypol reduces the amount of soluble protein and bioactive lysine in the meal (Zarins and Cherry, 1981).

Development of edible protein products from cotton seed for non-ruminants or humans has been impeded by the presence of the gossypol glands and has been a serious impediment for widespread cotton seed processing and use (Hopper, 1959). Mutant glandless cotton that is free of gossypol throughout the whole plant was first reported by McMichael (1954), and glandless cotton seed kernels have been used to produce snack foods, peanut butter, and baked and confectionary products. Despite initial optimism, these varieties have proven to be commercially unviable because the systemic absence of the protective terpenoids have made the aerial part of cotton plant much more susceptible to insect pests and pathogens than normal glanded cotton varieties.

Because (+)-δ-cadinene synthase catalyses the first committed step in gossypol biosynthesis, considerable efforts have also been made in characterizing and genetically manipulating this enzyme. A cDNA encoding (+)-δ-cadinene synthase was first cloned and functionally characterized from the A-genome diploid cotton (*G. arboreum*) (Chen *et al.,* 1995). The enzyme is encoded by a multigene family with different temporal and spatial regulation, and the isoforms may be responsible for different branches of the cotton sesquiterpene pathway. For example, (+)-δ-cadinene synthase mRNA and enzyme are highly up-regulated in response to infection by *Xanthomonas campestris* pv. *malvacearum* strains (now called *Xanthomonas axonopodis*) and *Verticillium dahliae* (Townsend *et al.,* 2005). RNAi down regulation of this gene led to a drastic seed-specific reduction of gossypol levels without reducing this compound and other related terpenoids in somatic tissues (Sunilkumar *et al.,* 2006). An average gossypol value of 0.2 µg/mg has been observed in the F_2 transgenic seeds, compared to 10 µg/mg in wild type, a value that falls well within the safety limit (0.6 µg/mg) set by World Health Organization (Lusas and Jividen, 1987).

Gene introgression from wild cotton species

An ideal cotton plant should have glandless seeds with low or no gossypol for optimal food and feed uses while retaining glanded foliage with a higher level of gossypol to limit pest attacks. Such a trait only exists naturally in some Australian wild diploid cotton species, such as *G. sturtianum* (Brubaker *et al.,* 1996). In cotton breeding programs inter-specific hybrids have been developed attempting to transfer the glandless-seed and glanded-plant trait from wild cotton to cultivated cotton (Vroh Bi *et al.,* 1999). A trispecies bridge hybrid strategy involving [(*G. hirsutum* x *G. raimondii*)^2 x *G. sturtianum*] was able to generate progeny plants that produced nearly or completely glandless seeds while still having normal gossypol gland density on their aerial parts. However, this combination of traits has yet to be introgressed into commercial cot-

ton cultivars because of the lack of recombination between the C genome of *G. sturtianum* and A and D genomes of tetraploid cotton. Attempts were also made to characterize the introgression of chromosomal segments from an Australian wild diploid cotton using RFLP markers of known chromosomal locations (Vroh Bi *et al.*, 1999). However, the mechanism for the repression of gossypol biosynthesis in the seeds of *G. sturtianum* remains unclear. It appears that multiple genes are involved as indicated by the high level of heterozygosity in the segregation of gossypol levels in the progeny testing following up to five generations of backcrossing and selfing (Benbouza *et al.*, 2010).

Germplasm selection for high (+)/(-) gossypol ratio

Gossypol is synthesized by a free radical dimerization of hemigossypol that yields two optically active enantiomers, (+)-gossypol and (-)-gossypol due to restricted rotation around the binaphthyl bond. Only (-)-gossypol is toxic to animals, while toxicity to insects and pathogens is independent of the (+)- to (-)- ratio. Therefore, a high (+)- to (-)- gossypol ratio in the seed would retain their natural defense capability while producing seeds that could be fed to non-ruminant animals. For example, broiler chickens fed a diet containing cotton seed with a higher ratio of (+)- to (-)- gossypol gained more weight compared to those fed on seeds with a low (+)- to (-)- gossypol ratio (Bailey *et al.*, 2000). On the other hand, a low ratio of (+)- to (-)-gossypol was found to be more effective in inhibiting the growth of various cancer cells, in anti-HIV activity, and in reducing male fertility (Lin *et al.*, 1993; Matlin *et al.*, 1985). The (+)- to (-)- gossypol ratio is genetically determined. In most US commercial cotton, the ratio is about 3:2. Several new sources of wild cottons that produce high levels of (+)-gossypol have been identified, including (+)- to (-)- gossypol ratio as high as 98:2 in some Moco cotton (*G. hirsutum* var. marie galante) accessions. Attempts have been made to breed cotton plants that incorporate such a trait from Moco cotton in order to maximise (+)-gossypol in the seed (Stipanovic *et al.*, 2000; Cai *et al.*, 2010).

Gossypol, especially in its (-)-enantiomer form, has great pharmacological interest due to its potential as an anti-cancer agent and for its male contraceptive effects (Dodou, 2005). Gossypol could effectively inhibit the enzyme activities of glycolysis and the TCA cycle, severely crippling energy metabolism and ATP production and lowering the mobility of human sperm cells in a dose-dependent manner (Medrano and Andreu, 1986).

CONCLUDING REMARKS

Currently, more than 1 billion people are undernourished with food intake below the recommended minimum daily energy requirement (OECD/FAO, 2010). Cotton seed, being a readily available by-product of more valuable cotton fiber production, is increasingly being recognized to have excellent potential as a source of additional food, feed and even biofuels for both developed and under-developed countries. In this review, we have attempted to emphasize the role of cotton seed as a renewable platform for the large-scale production and storage of many diverse biological molecules for agriculture and even for pharmaceutical and industrial enterprises. Value added cotton seed with broader applications through genetic improvement of both

seed production and quality without compromising fiber production is clearly in accord with the strategy for sustainable intensification of agricultural production by producing more food without increasing land use (Baulcombe, 2010).

Although histological, morphological, molecular and biochemical studies have provided descriptive information on embryogenesis and seed metabolism in cotton, the molecular and physiological events leading to the seed formation and storage compound accumulation are still far from being completely understood. Mutagenic silencing of genes and characterization of gene function through T-DNA tagging (techniques that are commonly used to generate loss-of-function alleles in model plants such as arabidopsis) might never be appropriate in cotton, primarily because of its relatively poor transformation efficiency and the inability to generate large numbers of transgenic lines. Furthermore, cotton is an allotetraploid with a moderately sized genome, and the presence of homeologous genes may mask knockout phenotypes because of the effects of redundant genes in the different genomes. However, the remarkable progress in the area of high throughput transcriptomics and proteomics over the last few years has now made large-scale investigations of genes and proteins achievable in cotton. Comprehensive web-based databases integrating updated EST collections, comparative transcriptome analyses between cotton and other plant species, expressed miRNAs and their putative targets, and simple sequence repeats (SSR) analyses, have become available (www.leonxi.com/, Xie *et al.,* 2011; cottonevolution.info/, Udall *et al*., 2006; and www.cotton marker.org/, Blenda *et al.*, 2006). In addition, because of its smaller genome size and low genome complexity, the worldwide cotton community has prioritized the D-genome progenitor *Gossypium raimondii* for complete sequencing by the Joint Genome Institute (Lin *et al.,* 2010) and a draft (as yet) unannotated assembly is now available for searching on-line (http://phytozome.net/cotton.php). Sequences of the A-genome progenitor *G. arboreum*, as well as a number of tetraploid (A+D) cottons are likely to follow in quick succession. These new genomic resources should now make it possible to study gene expression during cotton embryogenesis and seed development at a global level using whole genome microarrays or next generation sequencing technologies (RNAseq). They will no doubt provide an invaluable resource for identifying and characterizing genes that play critical roles during cotton embryogenesis and seed development. Insight into the complex process of seed development, and the identification and dynamic expression profiling during seed development of genes and proteins at the genome scale in many different plant species is beginning to provide a general framework for more in depth comparative studies that will aid our understanding of seed development and help identify targets for manipulation or transfer into crop plants to achieve specific desired outcomes.

Implementation of the tools of molecular biology and biotechnology has opened the door to the development of improved end-uses for cotton seed products for food, feed, as well as industrial applications. The biosynthetic pathways responsible for the synthesis and accumulation of many of the main storage compounds of seeds have largely been elucidated, and significant progress have already been made in the genetic improvement of a number of seed traits in cotton, including carbohydrate (Ruan *et al.,* 2003; Xu *et al.,* 2011), fatty acid composition (Chapman *et al.,* 2001; Liu *et al.,* 2002; 2008a,b) and gossypol (Sunilkumar *et al.,* 2006). The amenability of cotton seed to genetic modification has been clearly demonstrated in these ex-

periments, although none of the traits developed has yet to be commercialized. Despite this success, it is still unclear what factors determine the overall partitioning of seed reserves to the main storage components, and this will be critical if fiber development and quality is not to be impacted. More studies in cotton metabolic regulation will be necessary to understand the complex flux-control in the various biosynthetic pathways, especially in response to varying physiological and environmental conditions. Future advances in making a better cotton seed will be greatly facilitated by the wealth of genomics tools becoming available, and combined with efforts targeting metabolomics and flux map analysis, will allow for more rational design of genetic manipulation of the key metabolic control points to create substantial benefits for the global environment as well as the world-wide-economy and industry.

REFERENCES

Abid, G., J-M. Jacquemin, K. Sassi, Y. Muhovski, A. Toussaint, and J-P. Baudoin. 2010. Gene expression and genetic analysis during higher plants embryogenesis. Biotechnol. Agron. Soc. Environ. 14:667-680.

Alford, B.B., G.U. Liepa, and A.N. Vanbeber. 1996. Cotton seed protein: What does the future hold? Plant Foods for Hum. Nutr. 49:1-11.

Altenbach, S.B., C.C. Kuo, L.C. Staraci, K.W. Pearson, C. Wainwright, A. Georgescu, and J. Townsend. 1992. Accumulation of a Brazil nut albumin in seeds of transgenic canola results in enhanced levels of seed protein methionine. Plant Mol. Biol. 18:235-245.

Bailey, C.A., R.D. Stipanovic, M.S. Ziehr, A.U. Haq, M. Sattar, L.F. Kubena, H.L. Kim, and R.M. Vieira. 2000. Cottonseed with a high (+)- to (−)-gossypol enantiomer ratio favorable to broiler production. J. Agric. Food Chem. 48:5692-5695.

Bao, X., J.J. Thelen, G. Bonaventure, and J.B. Ohlrogge. 2003. Characterization of cyclopropane fatty-acid synthase from *Sterculia foetida*. J. Biol. Chem. 278:12846-12853.

Barozai, M.Y.K, M. Irfan, R. Yousaf, I. Ali, U. Qaisar, A. Maqbool, M. Zahoor, B. Rashid, T. Hussnain, and S. Riazuddin. 2008. Identification of microRNAs in cotton. Plant Physiol. Biochem. 46:739-751.

Baulcombe, D. 2010. Reaping benefits of crop research. Science 327:761.

Baumlein, H., S. Misera, H. LierBen, K. Kolle, C. Horstmann, U. Wobus, and A.J. Muller. 1994. The FUS3 gene of *Arabidopsis thaliana* is a regulator of gene expression during late embryogenesis. Plant J. 6:379-387.

Benbouza, H., J.M. Lacape, J.M. Jacquemin, B. Courtois, F.B.H. Diouf, D. Sarr, N. Konan, J.P. Baudoin, and G. Mergeai. 2010. Introgression of the low-gossypol seed & high-gossypol plant trait in upland cotton: Analysis of [(*Gossypium hirsutum* × *G. raimondii*)^2 × *G. sturtianum*] trispecific hybrid and selected derivatives using mapped SSRs. Mol. Breed. 25:273-286.

Benedict, C.K., R.J. Kohel, and A.M. Schubert. 1976. Transport of ^{14}C-assimilates to cottonseed: Integrity of funiculus during seed filling stage. Crop Sci. 16:23-21.

Berger, F., P.E. Grini, and A. Schnittger. 2006. Endosperm: an integrator of seed growth and development. Curr. Opin. Plant Biol. 9:664-670.

Blenda A., J. Scheffler, B. Scheffler, M. Palmer, J.M. Lacape, J. Yu, C. Jesudurai, S. Jung, S. Muthukumar, P. Yellambalase, S. Ficklin, M. Staton, R. Eschelman, M. Ulloa, S. Saha, B. Burr, S. Liu, T. Zhang, D. Fang, A. Pepper, S. Kumpatla, J. Jacobs, J. Tomkins, R. Cantrell, and D. Main. 2006. CMD: a cotton microsatellite database resource for *Gossypium* genomics. BMC Genomics 7:132.

Boothe, J., C. Nykiforuk, Y. Shen, S. Zaplachinski, S. Szarka, P. Kuhlman, E. Murray, D. Morck, and M.M. Moloney. 2010. Seed-based expression systems for plant molecular farming. Plant Biotechno. J. 8:588-606.

Bortman, S.J., R.N. Trelease, and J.A. Miernyk. 1981. Enzyme development and glyoxysome characterization in cotyledons of cotton seeds. Plant Physiol. 68:82-87.

Bourgis, F., A. Kilaru, X. Cao, G.F. Ngando-Ebongue, N. Drira, J.B. Ohlrogge, and V. Arondel. 2011. Comparative transcriptome and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning. Proc. Natl. Acad. Sci. USA 108:12527-12532.

Braybrook, S.A., and J.J. Harada. 2008. LECs go crazy in embryo development. Trends Plant Sci. 13:624-630.

Browse, J., and J.B. Ohlrogge. 1995. Lipid biosynthesis. Plant Cell 7:957-970.

Brubaker, C.L., C.G. Benson, and D.N. Miller. 1996. Occurrence of terpenoid aldehydes and lysigenous cavities in the "glandless" seeds of Australian *Gossypium* species. Aust. J. Bot. 44:601-612.

Cai, Y., Y. Xie, and J. Liu. 2010. Glandless seed and glanded plant research in cotton. A review. Agron. Sustain. Dev. 30:181-190.

Calhoun, M.C., S.W. Kuhlmann, and B.C. Baldwin Jr. 1995. Cotton feed product composition and gossypol availability and toxicity. pp. 125-145. *In*: Proceedings of the 2nd National Alternative Feeds Symposium. Alternative Feeds for Dairy and Beef Cattle, St. Louis, Mo.

Capdevila, A.M., and L. Dure III. 1977. Developmental biochemistry of cottonseed embryogenesis and germination. Plant Physiol. 59:268-273.

Casson, S.A., and K. Lindsey. 2006. The turnip mutant of Arabidopsis reveals that LEAFY COTYLEDON1 expression mediates the effects of auxin and sugars to promote embryonic cell identity. Plant Physiol. 142:526-541.

Cernac, A., and C. Benning. 2004. WRINKLED1 encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in Arabidopsis. Plant J. 40:575-585.

Chapman, K.D., J.M. Dyer, and R.T. Mullen. 2012. Biogenesis and functions of lipid droplets in plants: Thematic Review Series: Lipid Droplet Synthesis and Metabolism: from Yeast to Man. J. Lipid Res. 53:215-226.

Chapman, K.D., S. Austin-Brown, S.A. Sparace, A.J. Kinney, K.G. Ripp, I.L. Pirtle, and R.M. Pirtle. 2001. Transgenic cotton plants with increased seed oleic acid content. J. Am. Oil Chem. Soc. 78:941-947.

Chen, X-Y., Y. Chen, P. Heinstein, and V.J. Davisson. 1995. Cloning, expression, and characterization of (+)-*δ*-cadinene synthase: a catalyst for cotton phytoalexin biosynthesis. Arch. Biochem. Biophys. 324:255-266.

Cherry, J.P. 1983. Cottonseed oil. J. Am. Oil Chem. Soc. 60:360-367.

Cherry, J.P., and H.R. Leffler. 1984. p. 511-569. *In*: R.J. Kohel and C.F. Lewis (eds.). Cotton No. 24 in Agronomy series. ASA, CSSA, SSSA, Madison, Wisc.

Chlan, C.A., K. Borroto, J.A. Kamalay, and L. Dure III. 1987. Developmental biochemistry of cottonseed embryogenesis and germination. XIX. Sequence and genomic organization of the α globulin (vicilin) genes of cottonseed. Plant Mol. Biol. 9:533-546.

Clough, R.C., A.L. Mathis, S.R. Barnum, and J.G. Jaworsky. 1992. Purification and characterization of 3-ketoacyl-acyl carrier protein synthase III from spinach. A condensing enzyme utilizing acetyl-coenzyme A to initiate fatty acid biosynthesis. J. Biol. Chem. 267:20992-20998.

Coppock, C.E., J.K. Lanham, and J.L. Horner. 1987. A review of the nutritive value and utilization of whole cottonseed, cottonseed meal and associated by-products by dairy cattle. Anim. Feed Sci. Technol. 18:89-129.

Dahlqvist, A., U. Stahl, M. Lenman, A. Banas, M. Lee, L. Sandager, H. Ronne, and S. Stymne. 2000. Phospholipid:diacylglycerol acyltransferase: an enzyme that catalyses the acyl-CoA-independent formation of triacylglycerol in yeast and plants. Proc. Natl. Acad. Sci. USA 97:6487-6492.

Dewick, P.M. 2009. Medicinal natural products: A biosynthetic approach. 3rd ed. John Wiley & Sons, Ltd. West Sussex, United Kingdom. pp. 539.

Di Laurenzio, L., J. Wysocka-Diller, J.E. Malamy, L. Pysh, Y. Helariutta, G. Freshour, M.G. Hahn, K.A. Feldmann, and P.N. Benfey. 1996. The *SCARECROW* gene regulates an asymmetric cell division that is essential for generating the radial organization of the Arabidopsis root. Cell 86:423-433.

Dodeman, V., G. Ducreux, and M. Kreis. 1997. Zygotic embryogenesis versus somatic embryogenesis. J. Exp. Bot. 48:1493-1509.

Dodou, K. 2005. Investigations of gossypol - past and present developments. Expert Opin. Inv. Drugs 14:1419-1434.

Doman, D.C, J.C. Walker, R.N. Trelease, and B.N. Moore. 1982. Metabolism of carbohydrate and lipid reserves in germinated cotton seeds. Planta 155:502-510.

Dure III, L., and C. Chlan. 1981. Developmental biochemistry of cottonseed embryogenesis and germination. XII. Purification and properties of principal storage proteins. Plant Physiol. 68:180-186.

Dure III, L. 1975. Seed formation. Annu. Rev. Plant Physiol. 26:259-278.

Fischer, K., and A. Weber. 2002. Transport of carbon in non-green plastids. Trends Plant Sci. 7:345-351.

Fiume, E., and J.C. Fletcher. 2012. Regulation of Arabidopsis embryo and endosperm development by the polypeptide signalling molecule CLE8. Plant Cell 24:1000-1012.

Forman, M., and W.A. Jensen. 1965. Respiration and embryogenesis in cotton. Plant Physiol. 40:765-769.

Frizzi, A., S. Huang, L.A. Gilbertson, T.A. Armstrong, M.H. Luethy, and T.M. Malvar. 2008. Modifying lysine biosynthesis and catabolism in corn with a single bifunctional expression/silencing transgene cassette. Plant Biotechn. J. 6:13-21.

Fryxell, P.A. 1968. A redefinition of the tribe *Gossypieae*. Bot. Gaz. 129:296-308.

Galau, G.A., C.A. Chlan, and L. Dure III. 1983. Developmental biochemistry of cottonseed embryogenesis and germination. XVI. Analysis of the principal cotton storage protein gene family with cloned cDNA probes. Plant Mol. Biol. 2:189-198.

Galau, G.A., H.Y-C. Wang, and D.W. Hughes. 1992. Cotton *Mat5-A* (C164) gene and *Mat5-D* cDNAs encoding methionine-rich 2S albumin storage proteins. Plant Physiol. 99:779-782.

Goldberg, R.B., G. de Paiva, and R. Yadegari. 1994. Plant embryogenesis: zygote to seed. Science 266:605-614.

Haecker, A., R. Gross-Hardt, B. Geiges, A. Sarkar, H. Breuninger, M. Herrmann, and T. Laux. 2004. Expression dynamics of WOX genes mark cell fate decisions during early embryonic patterning in *Arabidopsis thaliana*. Development 131:657-668.

Haigler, C.H., M. Ivanova-Datcheva, P.S. Hogan, V.V. Salnikov, S. Hwang, K. Martin, and D.P. Delmer. 2001. Carbon partitioning to cellulose synthesis. Plant Mol. Biol. 47:29-51.

Hampson, S.E., R.S. Loomis, and D.W. Rains. 1979. Characteristics of sugar uptake in hypocotyls of cotton. Plant Physiol. 62:846-850.

Hendrix, D.L. 1990. Carbohydrates and carbohydrate enzymes in developing cotton ovules. Physiol. Plant. 78:85-92.

Herman, E.M., and A.W. Burks. 2011. The impact of plant biotechnology on food allergy. Cur. Opin. Biotechn. 22:224-230.

Hopper, T.H. 1959. Economic importance of gossypol. *In*: Proc. Conference on the Chemical Structure and Reactions of Gossypol and Nongossypol Pigments of Cottonseed, Southern Region Research Laboratory, USDA, New Orleans, LA. p. 12-13.

Hsu, S-C., M.F. Belmonte, J.J. Harada, and I. Kentaro. 2010. Indispensible role of plastids in *Arabidopsis thaliana* embryogenesis. Curr. Genomics 11:338-349.

Hu, G., N.L. Houston, D. Pathak, L. Schmidt, J.J. Thelen, and J.F. Wendel. 2011. Genomically biased accumulation of seed storage proteins in allopolyploid cotton. Genetics 189:1103-1115.

Huang, A.H.C. 1992. Oil bodies and oleosins in seeds. Annu. Rev. Plant Physiol. Plant Mol. Biol. 43:177-200.

Huang, S., A. Frizzi, C.A. Florida, D.E. Kruger, and M.H. Luethy. 2006. High lysine and high tryptophan transgenic maize resulting from the reduction of both 19- and 22-kD alpha-zeins. Plant Mol. Biol. 61:525-535.

Hughes, D.W., and G.A. Galau. 1989. Temporally modular gene expression during cotyledon development. Genes Dev. 3:358-369.

Hughes, D.W., H.Y-C. Wang, and G.A. Galau. 1993. Cotton (*Gossypium hirsutum*) MatP6 and MatP7 oleosin genes. Plant Physiol. (Plant Gene Register) 101:697-698.

Jako, C., A. Kumar, Y.D. Wei, J.T. Zou, D.L. Barton, E.M. Giblin, P.S. Covello, and D.C. Taylor. 2001. Seed-specific over-expression of an Arabidopsis cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight. Plant Physiol. 126:861-874.

Jenik, P.D., and M.K. Barton. 2005. Surge and destroy: the role of auxin in plant embryogenesis. Development 132:3577-3585.

Jolliffe, N.A., C.P. Craddock, and L. Fridgerio. 2005. Pathways for protein transport to seed storage vacuoles. Biochem. Soc. Trans. 33:1016-1018.

Kagaya, Y., R. Okuda, A. Ban, R. Toyoshima, K. Tsutsumida, H. Usui, A. Yamamoto, and T. Hattori. 2005. Indirect ABA-dependent regulation of seed storage protein genes by FUSCA3 transcription factor in Arabidopsis. Plant Cell Physiol. 46:300-311.

Kinney, A.J., R. Jung, and E.M. Herman. 2001. Cosuppression of the α'-subunits of β-conglycinin in transgenic soybean seeds induces the formation of endoplasmic reticulum-derived protein bodies. Plant Cell 13:1165-1178.

Kris-Etherton, P.M., V. Mustad, and J.A. Derr. 1993. Effects of dietary stearic acid on plasma lipids and thrombosis. Nutrition Today (USA). 28:30-38.

Lardizabal, K., R. Effertz, C. Levering, J. Mai, M.C. Pedroso, T. Jury, E. Aasen, K. Gruys, and K. Bennett. 2008. Expression of *Umbelopsis ramanniana* DGAT2A in seed increases oil in soybean. Plant Physiol. 148:89-96.

Li, W., L. Li, K. Li, J., J. Lin, X. Sun, X., and K. Tang. 2011. Expression of biologically active human insulin-like growth factor 1 in *Arabidopsis thaliana* seeds via oleosin fusion technology. Biotechnol. Appl. Bioc. 58:139-146.

Lin, L., G.J. Pierce, J.E. Bowers, J.C. Estill, R.O. Compton, L.K. Rainville, C. Kim, C. Lemke, J. Rong, H. Tang, X. Wang, M. Braidotti, A.H. Chen, K. Chicola, K. Collura, E. Epps, W. Golser, C. Grover, J. Ingles, S. Karunakaran, D. Kudrna, J. Olive, N. Tabassum, E. Um, M. Wissotski, Y. Yu, A. Zuccolo, M. Rahman, D.G. Peterson, R.A. Wing, J.F. Wendel, and A.H. Paterson. 2010. A draft physical map of a D-genome cotton species (*Gossypium raimondii*). BMC Genomics. 11:395.

Lin, T.S., R.F. Schinazi, J. Zhu, E. Birks, R. Carbone, Y. Si, K. Wu, L. Huang, and W.H. Prusoff. 1993. Anti-HIV-1 activity and cellular pharmacology of various analogs of gossypol. Biochem. Pharmacol. 46:251-255.

Liu, Q., S.P. Singh, K. Chapman, and A.G. Green. 2009. Bridging traditional and molecular genetics in modifying cottonseed oil. p. 353-384. *In:* A.H. Paterson (ed.). Plant Genetics and Genomics: Crops and Models Vol 3. Genetics and Genomics of Cotton. Springer, New York.

Liu, Q., P. Mansour, X-R. Zhou, J. Petrie, P.D. Nichols, and S. Singh, 2008a. Tailoring cottonseed oil for production of long-chain polyunsaturated fatty acids. *In*: Proc. 18th International Symposium on Plant Lipids. Bordeaux, France, 20-25 July 2008.

Liu, Q., S. Singh, P.J. Sharp, A.G. Green, and D.R. Marshall. 1996. Nucleotide sequence of a cDNA from *Gossypium hirsutum* encoding a stearoyl-acyl carrier protein desaturase (Accession No. X95988) (OGR96-012). Plant Physiol. 110:1435.

Liu, Q., S.P. Singh, and A.G. Green. 2002. High-stearic and high-oleic cottonseed oils produced by hairpin RNA-mediated post-transcriptional gene silencing. Plant Physiol. 129:1732-1743.

Liu, Q., S.P. Singh, and A. Green. 2008b. Genetic manipulation of fatty acid composition in cottonseed oil. *In:* 18th International Symposium on Plant Lipids. Bordeaux, France, 20-25 July 2008.

Liu, Q., S.P. Singh, C.L. Brubaker, and A.G. Green. 1999b. Cloning and sequence analysis of a novel member (accession No. Y10112) of the microsomal ω-6 fatty acid desaturase family from cotton (*Gossypium hirsutum*). Plant Physiol. 120:339.

Liu, Q., S.P. Singh, C.L. Brubaker, P.J. Sharp, A.G. Green, and D.R. Marshall. 1999a. Molecular cloning and expression of a cDNA encoding a microsomal ω-6 fatty acid desaturase in cotton (*Gossypium hirsutum* L.). Aust. J. Plant Physiol. 26:101-106.

Liu, Q., X. Bao, S. Singh, G. Pons, C. Hurlstone, C. Dowd, H. McFadden, M. Pollard, J.B. Ohlrogge, and A. Green. 2004. Studies on cotton cyclopropane fatty acid synthase. *In:* 16th International Symposium on Plant Lipids. Budapest, Hungary, 1-6 June 2004.

Long, J.A., E.I. Moan, J.I. Medford, and M.K. Barton. 1996. A member of the KNOTTED class of homeodomain proteins encoded by the STM gene of Arabidopsis. Nature 379:66-69.

Lotan, T., M. Ohto, K.M. Yee, M.A. West, R. Lo, R.W. Kwong, K. Yamagishi, R.L. Fischer, R.B. Goldberg, and J.J. Harada. 1998. Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. Cell 93:1195-1205.

Lowell, J.A., H.L. Parnes, and G.L. Blackburn. 1990. Dietary immunomodulation: Beneficial effects on oncogenesis and tumor growth. Critical Care Med. 18:5145-5148.

Lu, C., Z. Xin, Z. Ren, M. Miquel, and J. Browse. 2009. An enzyme regulating triacylglycerol composition is encoded by the *ROD1* gene of Arabidopsis. Proc. Natl. Acad. Sci. USA 106:18837-18842.

Lusas, E.W., and G.M. Jividen. 1987. Glandless cottonseed: a review of the first 25 years of processing and utilization research. J. Am. Oil Chem. Soc. 64:839-854.

Matlin, S.A., R. Zhou, G. Bialy, R.P. Blye, R.H. Naqvi, and M.C. Lindberg. 1985. (-)-Gossypol: an active male antifertility agent. Contraception 31:141-149.

McElver, J., I. Tzafrir, G. Aux, R. Rogers, C. Ashby, K. Smith, C. Thomas, A. Schetter, Q. Zhou, M.A. Cushman, J. Tossberg, T. Nickle, J.Z. Levin, M. Law, D. Meinke, and D. Patton. 2001. Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana*. Genetics. 159:1751-1763.

McMichael, S.C. 1954. Glandless boll in upland cotton and its use in the study of natural crossing. Agronomy J. 46:527-528.

Mozaffarian, D., M.B. Katan, A. Ascherio, M.J. Stampfer, and W.C. Willett. 2006. Trans fatty acids and cardiovascular disease. N. Engl. J. Med. 354:1601-1613.

Medrano, F.J., and J.M. Andreu. 1986. Binding of gossypol to purified tubulin and inhibition of its assembly into microtubules. Eur. J. Biochem. 158:63-69.

Mellon, J.E., and P.J. Cotty. 1999. Raffinose content may influence cottonseed susceptibility to aflatoxin contamination. J. Am. Oil Chem. Soc. 76:883-886.

Mensink, R.P., and M.B. Katan. 1992. Effect of dietary fatty acids on serum lipids and lipoproteins-a meta-analysis of 27 trials. Arterioscler. Thromb. 12:911-919.

Miernyk, J.A., and R.N. Trelease. 1981. Control of enzyme activities in cotton cotyledons during maturation and germination. IV. β-oxidation. Plant Physiol. 67:341-346.

Molvig, L., L.M. Tabe, B.O. Eggum, A.E. Moore, S. Craig, D. Spencer, and T.J. Higgins. 1997. Enhanced methionine levels and increased nutritive value of seeds of transgenic lupins (*Lupinus angustifolius* L.) expressing a sunflower seed albumin gene. Proc. Natl. Acad. Sci. USA 94:8393-8398.

Moncada, S., and A. Higgs. 1993. The L-Arginine-Nitric Oxide Pathway. N. Engl. J. Med. 329:2002-2012.

Muller, L.L., and T.J. Jacks. 1983. Intracellular distribution of free sugars in quiescent cottonseed. Plant Physiol. 71:703-704.

Murphy, C.M. 1993. Mass Spectrometry of Lipids. pp. 71-130. *In*: F. Snyder (ed.). Handbook of Lipid Research. Vol. 7. Plenum Press, New York.

Nodine, M.D., and D.P. Bartel. 2010. MicroRNAs prevent precocious gene expression and enable pattern formation during plant embryogenesis. Genes Dev. 24:2678-2692.

Nykiforuk, C.L., J.B. Boothe, E.W. Murray, R.G. Keon, J. Goren, N.A. Markley, and M.M. Moloney. 2005. Transgenic expression and recovery biologically active recombinant human insulin from *Arabidopsis thaliana* seeds. Plant Biotechno. J. 3:77-85.

Napier, J.A. 2007. The production of unusual fatty acids in transgenic plants. Annu. Rev. Plant Biol. 58:295-319.

O'Brien, R.D., L.A. Jones, C.C. King, P.J. Wakelyn, and P.J. Wan. 2005. Cotton seed oil. *In:* F. Shahidi (ed.). Bailey's Industrial Oil and Fat Products. 6th Edition. John Wiley & Sons, Inc. p. 173-279.

OECD/FAO. 2010. The growing problem. Nature 466:546-547.

Perry, H.J., R. Bligny, E. Gout, and J.L. Harwood. 1999. Changes in Kennedy pathway intermediates associated with increased triacylglycerol synthesis in oil-seed rape. Phytochemistry 52:799-804.

Pirtle, I.L., W. Kongcharoensuntorn, M. Nampaisansuk, J.E. Knesek, K.D. Chapman, and R.M. Pirtle. 2001. Molecular cloning and functional expression of the gene for a cotton Δ-12 fatty acid desaturase (FAD2). Biochim. Biophys. Acta 1522:122-129.

Price, H.J., and R.H. Smith. 1979. Somatic embryogenesis in suspension cultures of *Gossypium klotzschianum* Anderss. Planta 145:305-307.

Pugh, D.A., C.E. Offler, M.J. Talbot, and Y-L. Ruan. 2010. Evidence for the role of transfer cells in the evolutionary increase in seed and fiber biomass yield in cotton. Mol. Plant 3:1075-1086.

Reeves, R.G., and J.O. Beasley. 1935. The development of the cotton embryo. J. Agric. Res. 51:935-944.

Ruan, Y-L. 2005. Recent advances in understanding cotton fiber and seed development. Seed Sci. Res. 15:269-280.

Ruan, Y-L., D.J. Llewellyn, and R.T. Furbank. 2003. Suppression of sucrose synthase expression represses cotton fiber cell initiation, elongation and seed development. Plant Cell 15:952-964.

Ruan, R-L., D.J. Llewellyn, Q. Liu, S-M. Xu, L-M. Wu, L. Wang, and R.T. Furbank. 2008. Expression of sucrose synthase in developing endosperm is essential for early seed development. Func. Plant Biol. 35:382-393.

Sanjaya, T.P. Durrett, E. Sean, S.E. Weise, and C. Benning. 2011. Increasing the energy density of vegetative tissues by diverting carbon from starch to oil biosynthesis in transgenic Arabidopsis. Plant Biotechn. J. 9:874-883.

Santos-Mendoza, M., B. Dubreucq, S. Baud, F. Parcy, M. Caboche, and L. Lepiniec. 2008. Deciphering gene regulatory networks that control seed development and maturation in Arabidopsis. Plant J. 54:608-620.

Schubert, A.M., C.R. Benedict, and R.J. Kohel. 1986. Carbohydrate distribution in bolls. p. 311-324. *In:* J.R. Mauney and J.McD. Stewart (eds.). Cotton Physiology. The Cotton Foundation, Memphis, Tenn.

Schulz, P., and W.A. Jensen. 1977. Cotton embryogenesis: the early development of the free nuclear endosperm. Am. J. Botany 64:384-394.

Shen, B., W.B. Allen, P. Zheng, C. Li, K. Glassman, J. Ranch, D. Nubel, and M.C. Tarczynski. 2010. Expression of *ZmLEC1* and *ZmWRI1* increases seed oil production in maize. Plant Physiol. 153:980-987.

Shenstone, F.S., and J.R. Vickery. 1961. Occurrence of cyclo-propene acids in some plants of the order Malvales. Nature 190:168-169.

Shoemaker, R.C., I.J. Couche, and D.W. Galbraith. 1986. Characterization of somatic embryogenesis and plant regeneration in cotton (*Gossypium hirsutum* L.). Plant Cell Rep. 3:178-181.

Steinmann, T., N. Gelder, M. Grebe, S. Mangold, C.L. Jackson, S. Paris, L. Gälweiler, K. Palme, and G. Jürgens. 1999. Coordinated polar localization of auxin efflux carrier PIN1 by GNOM ARF GEF. Science 286:316-318.

Stewart, J.McD. 1986. Integrated events in the flower and fruit. pp. 261-300. *In*: J.R. Mauney and J.McD. Stewart (eds.). Cotton Physiology, Cotton Foundation, Memphis, Tenn.

Stipanovic, R.D., A.A. Bell, and M.J. Lukefahr. 1977. Natural insecticides from cotton (*Gossypium*). pp. 197-214. *In:* P.A. Hedin (ed.). Host Plant Resistance to Pests. Vol 62. American Chemical Society Symposium Series, Washington, DC.

Stipanovic, R.D., A.A. Bell, and C.R. Benedict. 1999. Cotton pest resistance: the role of pigment gland constituents. pp. 211-220. *In*: H.G. Cutler and S. Cutler (eds.). Biologically Active Natural Products: Agrochemicals. CRC Press. Boca Raton, Florida.

Stipanovic, R.D., C.A. Bailey, L.F. Kubena, and H.L. Kim. 2000. Cottonseed containing high levels of (+)-gossypol: Potential as a feed/food source. pp. 231-243. *In:* C. Benedict and G. Jividen, (eds.). Genetic Control of Cotton Fiber and Seed Quality; Cotton Incorporated, Cary, N.C.

Sunilkumar, G., J.P. Connell, C.W. Smith, A.S. Reddy, and K.S. Rathore. 2002. Cotton α-globulin promoter: isolation and functional characterization in transgenic cotton, Arabidopsis, and tobacco. Transgenic Res. 11:347-359.

Sunilkumar, G., L.M. Campbell, L. Puckhaber, R.D. Stipanovic, and K.S. Rathore. 2006. Engineering cottonseed for use in human nutrition by tissue-specific reduction of toxic gossypol. Proc. Natl. Acad. Sci. USA 103:18054-18059.

Szemenyei, H., M. Hannon, and J.A. Long. 2008. TOPLESS mediates auxin-dependent transcriptional repression during Arabidopsis embryogenesis. Science 319:1384-1386.

Taylor, D.C., Y. Zhang, A. Kumar, T. Francis, E.M. Giblin, D.L. Barton, J.R. Ferrie, A. Laroche, S. Shah, W. Zhu, C.L. Snyder, L. Hall, G. Rakow, J.L. Harwood, and R.J. Weselake. 2009. Molecular modification of triacylglycerol accumulation by over-expression of DGAT1 to produce canola with increased seed oil content under field conditions. Botany 87:533-543.

Townsend, B.J., A. Poole, C.J. Blake, and D.J. Llewellyn. 2005. Antisense suppression of a (+)-δ-cadinene synthase gene in cotton prevents the induction of this defense response gene during bacterial blight infection but not its constitutive expression. Plant Physiol. 138:516-528.

Trelease, R.N., J.A. Miernyk, J.S. Choinski Jr., and S.J. Bortman. 1986. Synthesis and compartmentalization of enzymes during cottonseed maturation. pp. 441-462. *In:* J.R. Mauney and J.McD. Stewart (eds.). Cotton Physiology. The Cotton Foundation, Memphis, Tenn.

Trolinder, N., and J.R. Goodin. 1987. Somatic embryogenesis and plant regeneration in cotton (*Gossypium hirsutum* L.). Plant Cell Rep. 6:231-234.

Trolinder, N., and J.R. Goodin. 1988. Somatic embryogenesis in cotton (*Gossypium*): II. Requirements for embryo development and plant regeneration. Plant Cell Tiss. Org. Cult. 12:43-53.

Trolinder, N.L., and C. Xhixian. 1989. Genotype specificity of the somatic embryogenesis response in cotton. Plant Cell Rep. 8:133-136.

Troncoso-Ponce, M.A., A. Kilaru, X. Cao, T.P. Durrett, J. Fan, J.K. Jensen, N.A. Thrower, M. Pauly, C. Wilkerson, and J.B. Ohlrogge. 2011. Comparative deep transcriptional profiling of four developing oilseeds. Plant J. 68:1014-1027.

Turley, R.B., and K.D. Chapman. 2010. Ontogeny of cotton seeds: gametogenesis, embryogenesis, germination, and seedling growth. pp. 332-341. *In:* J. McD. Stewart, D.M. Oosterhuis, J.J. Heiholt and J.R. Mauney, (eds.). Physiology of Cotton. Springer, New York.

Tzafrir, I., R. Pena-Muralla, A. Dickerman, M. Berg, R. Rogers, S. Hutchens, T.C. Sweeney, J. McElver, G. Aux, D. Patton, and D. Meinke. 2004. Identification of genes required for embryo development in Arabidopsis. Plant Physiol. 135:1206-1220.

Udall J.A., J.M. Swanson K. Haller, R.A. Rapp, M.E. Sparks, J. Hatfield, Y. Yu, Y. Wu, C. Dowd, A.B. Arpat, B.A. Sickler, T.A. Wilkins, J.Y. Guo, X.Y. Chen, J. Scheffler, E. Taliercio, R. Turley, H. McFadden, P. Payton, N. Klueva, R. Allen, D. Zhang, C. Haigler, C. Wilkerson, J. Suo, S.R. Schulze, M.L. Pierce, M. Essenberg, H. Kim, D.J. Llewellyn, E.S. Dennis, D. Kudrna, R. Wing, A.H. Paterson, C. Soderlund, and J.F. Wendel. 2006. A global assembly of cotton ESTs. Genome Res. 16:441-450.

Voinnet, O. 2009. Origin, biogenesis, and activity of plant microRNAs. Cell 136:669-687.

Vroh Bi, I., A. Maquet, J.P. Baudoin, P. Jardi, J.M. Jacquemin, and G. Mergeai. 1999. Breeding for "low-gossypol seed and high-gossypol plants" in upland cotton, Analysis of tri-species hybrids and backcross progenies using AFLPs and mapped RFLPs. Theor. Appl. Genet. 99:1233-1244.

Wang, M., Q. Wang, B., and Wang. 2012. Identification and characterisation of microRNAs in Asiatic cotton (*Gossypium arboreum* L.). PLoS ONE 7:e33696.

Weselake, R.J., S. Shah, M. Tang, P.A. Quant, C.L. Snyder, T.L. Furukawa-Stoffer, W. Zhu, D.C. Taylor, J. Zou, A. Kumar, L. Hall, A. Laroche, G. Rakow, P. Raney, M.M. Moloney, and J.L. Harwood. 2008. Metabolic control analysis is helpful for informed genetic manipulation of oilseed rape (*Brassica napus*) to increase seed oil content. J. Exp. Bot. 59:3543-3549.

West, M.A.L., and J.J. Harada. 1993. Embryogenesis in higher plants: An overview. Plant Cell 5:1361-1369.

Willemsen, V., and B. Scheres. 2004. Mechanisms of pattern formation in plant embryogenesis. Annu. Rev. Genet. 38:587-614.

Willmann, M.R., A.J. Mehalick, R.L. Packer, and P.D. Jenik. 2011. MicroRNAs regulate the timing of embryo maturation in Arabidopsis. Plant Physiol. 155:1871-1884.

Wood, E.M., and W.T. Yasutake. 1956. Ceroid in fish. Am. J. Pathology 32:591-603.

Wood, R. 1986. Comparison of the cyclopropene fatty acid content of cottonseed varieties, glanded and glandless seeds, and various seed structures. Biochem. Arch. 2:73-80.

Wu, X., F. Li, C. Zhang, C. Liu, and X. Zhang. 2009. Differential gene expression of cotton cultivar CCRI24 during somatic embryogenesis. J. Plant Physiol. 166:1275-1283.

Wysocka-Diller, J.W., Y. Helariutta, H. Fukaki, J.E. Malamy, and P.N. Benfey. 2000. Molecular analysis of SCARECROW function reveals a radial patterning mechanism common to root and shoot. Development 127:595-603.

Xiao, W., K.D. Custard, R.C. Brown, B.E. Lemmon, J.J. Harada, R.B. Goldberg, and R.L. Fischer. 2006. DNA methylation is critical for Arabidopsis embryogenesis and seed viability. Plant Cell 18:805-814.

Xie, F., G. Sun, J.W. Stiller, and B. Zhang. 2011. Genome-wide functional analysis of the cotton transcriptome by creating an integrated EST database. PLoS ONE 6:e26980.

Xu, S-M., E. Brill, D.J. Llewellyn, R.T. Furbank, and Y-L. Ruan. 2011. Overexpression of a potato sucrose synthase gene in cotton accelerates leaf expansion, reduces seed abortion, and enhances fiber production. Mol. Plant 5:430-441.

Yamagishi, K., N. Nagata, K.M. Yee, S.A. Braybrook, J. Pelletier, S. Fujioka, S. Yoshida, R.L. Fischer, R.B. Goldberg, and J.J. Harada. 2005. TANMEI/EMB2757 encodes a WD repeat protein required for embryo development in Arabidopsis. Plant Physiol. 139:163-173.

Youle, R.J., and A.H.C. Huang. 1979. Albumin storage protein and allergens in cottonseeds. J. Agri. Food Chem. 27:500-503.

Youle, R.J., and A.H.C. Huang. 1981. Occurrence of low molecular weight and high cysteine containing albumin storage proteins in oilseeds of diverse species. Amer. J. Bot. 68:5.

Yu, X-H., R. Rawat, and J. Shanklin. 2011. Characterization and analysis of the cotton cyclopropane fatty acid synthase family and their contribution to cyclopropane fatty acid synthesis. BMC Plant Biol. 11:97-106.

Zarins, Z.M., and J.P. Cherry. 1981. Storage proteins of glandless cottonseed flour. J. Food Sci. 46:1855-1859.

Zeng, F., X. Zhang, L. Zhu, L. Tu, X. Guo, and Y. Nie. 2006. Isolation and characterization of genes associated to cotton somatic embryogenesis by suppression subtractive hybridization and macroarray. Plant Mol. Biol. 60:167-183.

Zeng, F., X. Zhang, L. Cheng, L. Hu, L. Zhu, J. Cao, and X. Guo. 2007. A draft gene regulatory network for cellular totipotency reprogramming during plant somatic embryogenesis. Genomics 90:620-628.

Zhang, B., Q. Wang, K. Wang, X. Pan, F. Liu, T. Guo, G.P. Cobb, and T.A. Anderson. 2007. Identification of cotton microRNAs and their targets. Gene 397:26-37.

Zhang, M, J.L. Fan, D.C. Taylor, and J.B. Ohlrogge. 2009. DGAT1 and PDAT1 acyltransferases have overlapping functions in Arabidopsis triacylglycerol biosynthesis and are essential for normal pollen and seed development. Plant Cell 21:3885-3901.

Zhang, D., I.L. Pirtle, S.J. Park, M. Nampaisansuk, P. Neogi, S.W. Wanjie, R.M. Pirtle, and K.D. Chapman. 2010. Identification and expression of a new delta-12 fatty acid desaturase (FAD2-4) gene in upland cotton and its functional expression in yeast and *Arabidopsis thaliana* plants. Plant Physiol. Bioch. 47:462-471.

Zheng, P., W.B. Allen, K. Roesler, M.E. Williams, S. Zhang, J. Li, K. Glassman, J. Ranch, D. Nubel, W. Solawetz, D. Bhattramakki, V. Llaca, S. Deschamps, G.Y. Zhong, M.C. Tarczynski, and B. Shen. 2008. A phenylalanine in DGAT is a key determinant of oil content and composition in maize. Nat. Genet. 40:367-372.

Zhou, X-R., S. Singh, Q. Liu, and A. Green. 2006. Combined transgenic expression of Δ12-desaturase and Δ12-epoxygenase in high linoleic acid seeds to increased accumulation of vernolic acid. Func. Plant Biol. 33:585-592.