Genetic contributions to agricultural sustainability

Elizabeth S. Dennis, Jeffrey Ellis, Allan Green, Danny Llewellyn, Matthew Morell, Linda Tabe and W. J. Peacock*

CSIRO Plant Industry, GPO Box 1600, Canberra, Australian Capital Territory 2601, Australia

The current tools of enquiry into the structure and operation of the plant genome have provided us with an understanding of plant development and function far beyond the state of knowledge that we had previously. We know about key genetic controls repressing or stimulating the cascades of gene expression that move a plant through stages in its life cycle, facilitating the morphogenesis of vegetative and reproductive tissues and organs. The new technologies are enabling the identification of key gene activity responses to the range of biotic and abiotic challenges experienced by plants. In the past, plant breeders produced new varieties with changes in the phases of development, modifications of plant architecture and improved levels of tolerance and resistance to environmental and biotic challenges by identifying the required phenotypes in a few plants among the large numbers of plants in a breeding population. Now our increased knowledge and powerful gene sequence-based diagnostics provide plant breeders with more precise selection objectives and assays to operate in rationally planned crop improvement programmes. We can expect yield potential to increase and harvested product quality portfolios to better fit an increasing diversity of market requirements. The new genetics will connect agriculture to sectors beyond the food, feed and fibre industries; agri-business will contribute to public health and will provide high-value products to the pharmaceutical industry as well as to industries previously based on petroleum feedstocks and chemical modification processes.

Keywords: genomics; transgenics; human nutrition; sustainable production; plant breeding

1. INTRODUCTION

Yields in most agricultural crops have been increasing steadily over the past 100 years. In most cases, advances in both agronomic management and plant improvement programmes have contributed significantly to these increases. In recent years, these two components of yield improvement have become more intimately intertwined with inbuilt genetic traits delivered in the seed being able to replace some management inputs, particularly in pest control. Improvements in management have closed the gap between best farm yield and yield potential of the crop for a range of input regimes. In parallel, the average farm yields have approached best farm yields as a consequence of better extension services, accessible computer decision support tools and increased abilities of farmers to recognize and adopt best industry practice.

Yield potential has not reached an asymptote even in the most extensively improved crops such as maize, wheat, rice, soya bean and cotton. The biological potential of these crops has continued to be increased by plant breeding systems aimed at increasing harvest index, water use efficiency, nutrient acquisition and genetic protection against biotic and abiotic challenges. Protection against pathogens and pests often enables a crop to continue to be produced in an area where the entry of a virulent pathogen or the evolution of a new strain of an endemic pathogen would otherwise have made production uneconomic. For example, rust devastated wheat yields in regions of Australia until breeders produced varieties resistant to the pathogen challenge. There are similar examples for all crops and their accompanying pathogens and pests.

Plant breeders have been remarkably skilled in identifying sources of resistance genes or, more correctly, resistance alleles in wild relatives of crop species and in the extensive germplasm collections available for most major crops that can act as gene donors through sexual reproductive methods. Embryo rescue and genetic system manipulations are frequently needed to access alleles from distantly related species. Rarely the introduced genes have been unrelated to the genes of the crop species and, in general, they are usually accommodated by the metabolic and cellular pathways already existing in the crop species.

Not all breeding goals have been met by the introduction of single genes in different allelic forms. In many cases, the breeder has had to cope with introducing alleles of several loci, often unlinked, which have products that interact to produce the desired phenotype. Breeders have also faced another hurdle of gene product interaction where the entry of the new allele produces the needed phenotype, but has pleiotropic, sometimes negative, effects on other traits. In the past, these cases of polygenic inheritance with pleiotropic effects have been dealt with by various strategies in the construction of breeding systems, mostly without any obvious understanding of what is happening at the molecular, cellular or tissue levels of plant function. These complications
have caused major obstacles in achieving breeding system objectives.

Recent advances in our understanding of how plants function and develop have increased the power and efficiency of plant improvement programmes. The knowledge of gene and genome sequences, the regulation of gene expression and the molecular and cellular mechanisms and pathways behind plant architecture, development and function, have provided new opportunities for breeders to rationally design improvement programmes providing for more homeostasis in the environmental responses of a crop and to better mould the phases and components of plant development within the constraints set by the crop environment.

Breeders are confronted with difficulties in their selection programmes owing to heterogeneity in field-based bioassays and the uncertainty of environmental pressures.

Knowledge of DNA sequence has provided sequence markers for desired alleles and, in some cases, these enable a breeder to bypass bioassays and environmental assays that previously had introduced major constraints and unreliability into breeding programmes. Markers for different genetic traits have also facilitated the stacking of duplicate systems of protection or function thus providing for a more robust and stable phenotype. Gene interactions and pleiotropic effects are now frequently understood at a molecular and cellular level with breeders being able to specifically avoid some of the negative interactions, e.g. selection for a subset of functions of certain transcription factors or selection for more specific phenotypes with a high level of understanding of the feedback loops in metabolic pathways and their impacts on phenotype.

The need for more powerful and more efficient plant breeding for our major food plants is, in part, driven by the rate of global population increase. There is a need to produce more food but there is no more arable land available, so the food supply needed over the next 30–40 years, approximately twice what we are producing now, has to be produced on existing land. In fact, since a significant proportion of the lands on which we now have agricultural production is already marginal, with continued agriculture leading to increasing damage, then the challenge is to produce the required additional food on rather less land than we use now.

In addition, with the increased knowledge of the way in which ecosystems are working in our world, our societies are demanding that our agricultural production systems work with empathy to the environment and not in opposition. With the inevitable removal of nutritional factors from the soil, there is a realization that these components must be replaced on a regular basis and, at the same time, the agricultural production lands should not be seen to be damaging the adjacent non-agricultural production ecosystems. In addition, the basic natural resources of the water supply systems, the soil and the quality of the air should not be put under threat as a consequence of agricultural activities. If we can achieve all of these requirements, then we will approach the sustainability needed for the food production systems of the world.

Apart from the quantity of food to be produced, our increased knowledge of human nutrition has highlighted that many plant food commodities and processed food products are less than optimal for human health. There are clear deficiencies in our food supply in both developed and developing countries as evidenced by major non-infectious and non-communicable diseases, such as heart disease and diabetes. In developing countries, nutritional deficiencies are particularly evident when one major food commodity forms a major staple of the diet. In more developed countries where there is ample choice for food, the fault lies with both dietary habits and lifestyle choices.

With the increased knowledge of human nutritional requirements and our increasing abilities to modify the characteristics of our food plants, we have a clear expectation that modern plant breeding should be able to enhance the nutritional qualities of our major food plants so that they approach the optimal composition for human health regardless of lifestyle. There is ample evidence that in many cases the frequency and severity of non-communicable diseases can be reduced with a suitable diet. In the case of type 2 diabetes, which is rapidly increasing in both developing and developed countries, medical studies have shown that the consumption of reduced glycaemic index foods can decrease its incidence. Glycaemic index is primarily a property of the starchy components of our major cereal food plants. In crops such as rice, the deficient characteristics of the major component of the grain, are mirrored in the other grain components such as protein, fatty acids and sugars that all have less than optimal make-up. There are considerable opportunities for plant breeding to address these deficiencies.

Plant breeding can also address other major societal objectives such as the increasing global energy requirement and the depletion of petroleum-based resources. Bio-based, and hence renewable, energy sources are attracting greater attention around the world. Apart from the production of ethanol from fermented plant sugars, methane from waste plant products and biodiesel from plant oils, it is highly probable that crops will also be able to produce pharmaceutical and other valuable industrial products with the advantage of a sustainable supply chain based on agriculture.

The prospects of achieving these alterations to our crop plants have been enormously increased with the advent of the new technologies of genetic modification and new levels of knowledge of plant genomes. There is still a major problem to tackle with these new developments, in that many societies of the world are not willing, at this time, to accept food and other products derived from transgenic crops, crops in which some key traits have been provided through laboratory technologies and not through sexual reproductive systems. Communities in most countries around the world now readily accept genetically engineered products in medicine, such as genetically engineered human insulin and human growth hormone, largely because the products can be very easily seen by consumers to improve their quality of life. Genetically modified (GM) crops, on the other hand, have received a lower level of acceptance and for many people there is even downright fear and antagonism. Many are concerned that GM crops mostly improve the profitability of big business and not the quality of life for...
the individual consumer. This view is likely to change with time as the first generation of GM crops is seen to deliver both economic and environmental benefits to farmers and societies without the devastating ecological or health impacts predicted by their detractors and as the technologies deliver new crops that produce healthier foods, or that produce drugs and vaccines providing simple delivery mechanisms to improve health in developing countries.

Another reason for the poor public image of new GM crops is the misinformation that has been so frequently presented in the media around the world. Scientists have been successful using the new technologies in achieving remarkable advances in the knowledge of how plants develop and function. We have vastly increased capacities to provide appropriate genetic instructions in our major crop plants so that they perform optimally in a range of environments and provide us with core food products tailored to human nutritional requirements. But, as scientists, we have failed to effectively combat the campaigns of misinformation that have been pedalled by various activist groups who, for one reason or another, speak ill of the introduction of the much needed improvements to our food production systems.

In this paper, we provide a number of examples showing how the improved plant breeding capacities engendered by the new developments in genetics, genomics and genetic modification are likely to enhance the performance and quality of the products from our crop plants. The first generation of GM crops has improved the efficiency and economy of production through enhanced pest and weed control. In many cases, the genetic changes promised by the new era of genomic information will deliver improvements that will be applicable to the end product consumer rather than mostly to the producer, including the production of crops with novel oils and starches with improved nutritional benefits or other industrial non-food uses. As more of these second generation products reach the market, so too will public acceptance of transgenic food products increase to the point where society can reap the huge health and agricultural sustainability benefits that these new technologies can deliver.

2. THE FUNCTION AND REGULATION OF PLANT GENES—GENOME-WIDE ANALYSES PROVIDING A FIRM FOUNDATION FOR THE NEW GENETICS IN CROP IMPROVEMENT

The ways in which plants develop and respond to the environment in order to produce an optimal yield of food or fibre is the result of the controlled expression of the approximately 30 000 genes that are present in the genome of all plants. The role of genomics is to define the function of these genes, determine how they are regulated and how their gene products interact. These findings can then be applied to crop improvement.

The genome of Arabidopsis thaliana, a dicot (or broad-leaved species) related to the Brassicas (like canola and cabbage), and the genome of rice, a monocot, have been completely sequenced (The Arabidopsis Genome Initiative 2000; Goff et al. 2002). Sequencing of the genomes of a number of other species—maize, lotus, Medicago truncatula, poplar, grape and tomato—is underway (www.ncbi.nlm.nih.gov/genomics). The two sequenced genomes serve as basic references for all crop plants—Arabidopsis for the dicots and rice for the cereals. The gene content of these two widely divergent species is remarkably similar and it is probable that further genome sequencing will reinforce the similarity of gene content across all the flowering plants. The similarity of gene make-up of the genome of different species is not necessarily mirrored by the way in which these genes are regulated or by the interaction of their products in regulatory networks; these properties can differ markedly. It is this difference in patterns of gene expression that differentiate species. Genes that specify secondary metabolites or particular attributes, such as structural properties, have evolved from a common pool of genes. Gene duplication, such as occurs in polyploidy, and acquisition of separate functions or separate patterns of expression by each of the duplicated genes has been a frequent avenue for providing variation to be acted on by natural selection (Adams & Wendel 2005).

Genomics can assist in identifying which genes are involved in specifying particular characteristics of a plant. The first step in identifying the function of a gene is to compare its nucleotide or amino acid sequence with all of the sequences in databases derived from the genomes of other organisms. A function may be assigned through similarity to other genes with known function, hence genomes can have usefulness across species or even across kingdoms in allowing us to specify function. Genome-wide mutagenesis using transposable elements such as Ac/Ds, Tos17 (an endogenous retrotransposon of rice) or T-DNA insertions has resulted in the production of populations consisting of many lines, where each line contains an insert in a single gene. Since the DNA sequence of the insert is known, it is simple to determine which gene has been disrupted by cloning the flanking sequence. There is a set of Arabidopsis lines containing inserts in approximately 80% of the genes and, in rice, a similar proportion are tagged; these lines are freely available (www.arabidopsis.org/abrc/ecker_frank.jsp and Hirochika et al. 2004). These tagged lines can be made homozygous and their phenotypes determined to associate a gene with a specific phenotype. The tagged genes can then become candidates for crop improvement either as DNA markers or directly in transgenic breeding.

Interruption of gene activity can also be generated by RNAi, a supplementary form of mutagenesis, by which a construct introduced into a plant gives rise to a double-stranded RNA that activates a sequence-specific degradation mechanism that disrupts the mRNA of the gene target which may produce a phenotype (Wang & Waterhouse 2002). The advantages of RNAi for functional genomics are that RNA constructs targeted to a gene act in a dominant manner, and a gene in any background (e.g. a mutant background) can be targeted. Recently, synthetic microRNAs have also been used for gene silencing. MicroRNAs play a role in the control of genes involved in plant development and stress response so the new technology provides an additional option for silencing. Because microRNAs are shorter (21 nt) compared with the 200–300 bp usually targeted by RNAi, conserved
regions in gene families can be targeted, silencing multiple genes simultaneously (Alvarez et al. 2006; Schwab et al. 2006).

An essential aspect of applying genomics to crop improvement is that there must be the ability to use high-throughput technologies to screen for changes in phenotype. Phenotyping can involve automated growth measurements and imaging under various environmental stress conditions. It should also involve field-based screening as characters that appear useful in the glasshouse are sometimes not maintained in the field. Such high-throughput strategies have formed the basis for a number of international consortia to characterize mutations or silenced lines in specific classes of Arabidopsis genes (e.g. Agrikola, to identify the functions of specific types of transcription factor genes (Hilson et al. 2004)).

A second resource that complements and extends genome sequences is gene arrays (microarrays), which consist of large numbers of oligonucleotides or cDNAs arrayed on slides. For the sequenced genomes all predicted genes can be included on the arrays. The arrays can then be hybridized to RNA extracted from a particular tissue or developmental stage from a mutant, or from a plant subjected to environmental or disease stress to determine which genes are preferentially expressed compared with expression in control wild-type plants. The basic principle underlying microarrays is that if a gene is expressed under certain conditions, it may play a role in that condition, for example, genes induced by salt may provide protection in saline conditions. For genomes that are not sequenced, expressed sequence tags (ESTs) or anonymous cDNAs can be arrayed and hybridized in the same way and candidate genes sequenced later. Using microarrays, genes with expression patterns similar to those of known genes can be chosen giving a greater choice of target, e.g. under anaerobic conditions genes can be chosen with a similar expression response pattern to alcohol dehydrogenase suggesting a similar involvement in the anaerobic response. Data from many thousands of Arabidopsis microarray experiments have been gathered together in the Genevestigator database (www.genevestigator.ethz.ch/) where data from different mutants, different stages of development or different stresses are all gathered together for public access. A reduction in the cost of sequencing has also fostered the development of high-throughput sequencing of expressed genes in cDNA libraries as an indicator of gene expression levels. Only short stretches of sequence are required to match an EST to its gene sequence, so massively parallel signature sequencing strategies provide a good starting point for determining the expression levels of a particular gene in the particular tissues or conditions of treatment of the plants from which the libraries are made. These tools provide a ready knowledge of where and when any particular gene is expressed. The conservation of many genes and biological processes between species means that expression patterns are also likely to be conserved, thus interrogating the Arabidopsis expression databases with a gene sequence from another plant can also provide some useful information in defining the functional roles of that gene (e.g. Zhang et al. 2004).

Once a candidate gene has been identified from sequence comparisons and its expression patterns, the experimental increasing or decreasing of its level of activity (constitutive or tissue-specific overexpression of the gene or complete or partial inactivation of the gene by insertion mutants or RNAi) can be used to confirm its importance in a specific gene pathway. Knockout and overexpression lines of a specific gene can then be passaged back through a microarray experiment, for example, to determine which other genes may be affected by its expression or lack of expression helping to complete our knowledge of some of the regulatory networks that exist in plants.

The availability of complete genome sequences allows the production of tiling arrays in which all the bases in the genome are arrayed in an overlapping manner, not just the coding regions of the genes. Tiling arrays allow coding region transcripts to be assayed as well as transcripts which are not associated with coding regions such as small regulatory RNAs, now known to play an important role in gene regulation (www.affymetrix.com/products/arrays/specific/arab_tiling.affx). Whole-genome arrays can also be used for determining changes in gene expression determined by epigenetic mechanisms such as DNA methylation or histone modification. The epigenetic controls of gene activity can be further probed by techniques such as immunoprecipitation of proteins bound to DNA, which may be transcription factors or other regulatory proteins. This allows the mapping of regulatory regions in the genome and identifies which transcription factors regulate different genes. The genome arrays also assist in rapid sequencing of related species or cultivars of the same species to help determine the association between sequences and phenotypes, defining linkages or gene structures conserved in speciation and evolution.

All of these new tools are enabling a greater understanding of the genes controlling the physiological processes of crop plants. For example, in cotton, genes expressed at the early stages of fibre initiation in plants that produce fibre were compared with those expressed in fibreless mutants. Genes that were not expressed in the mutants were identified. These genes may play an important part in fibre initiation and quality (Wu et al. 2006). The effect of altering the levels of these genes must then be tested to confirm their role in fibre formation and whether they will have use in producing cotton plants with improved fibre yield or fibre quality.

During early seed development, the ability to monitor changes in gene expression in very small samples has demonstrated that many, but not all, genes expressed in the developing seed show differential expression dependent upon whether they are derived from the maternal or paternal genome, i.e. genes are imprinted. This imprinting control of seed-expressed genes is one example of epigenetic regulation, that is, control by changes in the architecture of the DNA rather than its sequence (Autran et al. 2005). Epigenetic control of gene expression can be mediated through repressive protein complexes (e.g. polycomb complexes) that carry out chromatin modification and affect gene expression. The protein components of polycomb complexes are similar in plants, Drosophila, mouse and humans, implying a conserved process for
this mechanism of epigenetic control. In plants, many important developmental transitions are controlled by polycomb group proteins including vernalization, the transition to flowering and seed development (Kohler & Grossniklaus 2002). The protein components involved in complexes controlling various other developmental processes are being identified based on genomic approaches.

In many cases, decreased activity of a gene will be needed for crop improvement. Gene silencing using the RNAi technology is now being used widely to achieve novel phenotypes in crop plants (see §§ 4 & 5), but concerns over the use of genetic modification are limiting the delivery of such new traits beyond the laboratory. With the genome sequence and the definition of specific candidate genes for any desired character, it is now possible to produce new mutant alleles of a locus by targeting induced local lesions in genomics (TILLING; Henikoff et al. 2004)—a non-GM method. This will allow researchers to reach a desired phenotype without using transgenic plants. In TILLING, the plant genome is subjected to extensive mutagenesis; polymerase chain reaction (PCR) is then used to amplify the candidate gene or promoter and sensitive methods used to detect any base changes within this target sequence. A similar technique, EcoTILLING (Comai et al. 2004) can be used to identify naturally occurring variants that can then be incorporated into breeding programmes.

The excitement and the challenge in controlling gene expression for plant improvement programmes is that gene expression is exquisitely sensitive to many factors, developmental and environmental, yet genes do not work alone. Genes interact in complex networks so changes in the expression of a single gene can have dramatic effects in multiple pathways. Understanding and modelling the interactions involved in gene regulatory networks is a new goal for genomics. Success will lead to a greatly enhanced ability to harness gene activities for plant improvement and this new genetic knowledge will undoubtedly underpin the next generation of improved crop plants. Scientific developments in the understanding of these processes have outstripped their application in breeding programmes. In the interim, we have well-established methods for modifying the expression of small numbers of genes using transgenic plants, mutant screening or combining natural alleles and these are starting to contribute new traits such as those described below.

3. IMPROVING THE ESSENTIAL AMINO ACID BALANCE IN PLANT PROTEINS USED FOR FOOD AND FEED

Seeds are major sources of dietary protein for large vegetarian populations around the world and intensively farmed animals. However, the protein in seeds can have a skewed amino acid composition due to the high abundance of a limited number of individual seed storage proteins. Of the 20 protein amino acids, 10 are classified as ‘essential’ because they cannot be synthesized by animals, and consequently must be obtained from the diet. Insufficiency of certain essential amino acids can be a cause of malnutrition in countries that are dependent on a diet of low diversity and can limit the efficiency of animal production. Legume and cereal grains are particularly important for human and animal nutrition, but their seed protein is deficient in the essential amino acids methionine and lysine, respectively (Tabe & Higgins 1998; Amir & Galili 2003). These deficiencies can be offset to some extent by combining the two types of seeds, but animal feeds are still supplemented with synthetic amino acids for optimal nutrition (Habben & Larkins 1995). In developing countries, up to 90% of food intake can be derived from a single crop species, so amino acid balance of individual seeds becomes a critical consideration also for human nutrition.

In recent years, both genetic modification and plant breeding with induced or natural mutants have achieved important successes in modifying amino acid composition of cereals and legumes. This section is focused on modification of mainly grain legumes to improve their content of the essential, sulphur-containing amino acid, methionine. Three approaches have been used: genetic modification to increase methionine biosynthesis; genetic modification to increase methionine storage in protein; and selection of mutants with increased methionine.

(a) Engineering the methionine biosynthetic pathway in plants

Sulphur is taken up from the soil in the oxidized form of sulphate and is subsequently reduced in the plastids of plant cells, then incorporated into an amino acid backbone derived from serine via the action of the enzyme serine acetyltransferase. The product of this reaction is cysteine, the first stable reduced sulphur metabolite in the cell, and a substrate for many other biochemical pathways. Methionine is derived from cysteine by the sequential action of three enzymes, the first of which, cystathionine γ-synthase (CGS), combines O-phosphohomoserine from the aspartate amino acid pathway and cysteine (Leustek & Saito 1999). There are numerous reports in the literature of genetic manipulation of the activities of the enzymes of the reductive sulphur assimilation and sulphur amino acid biosynthesis (reviewed by Amir & Tabe (2006)). Some dramatic increases in free cysteine and methionine have been observed in the leaves of the GM plants, sometimes at specific growth stages. However, free amino acids are much less abundant in planta than protein-bound amino acids. Consequently, in the few cases where total amino acid composition was analysed, these manipulations had relatively minor effects on total methionine concentration. For example, constitutive expression of a CGS enzyme from A. thaliana in GM tobacco or GM alfalfa increased free methionine in the leaves, but had no significant effect on protein-bound methionine (Hacham et al. 2002; Bagga et al. 2005). On the other hand, in a rare exception to this generalization, expression of a mutated form of CGS in GM tobacco resulted in not only a large increase in free methionine in the leaves, but also a twofold increase in protein-bound methionine compared with controls. The high-methionine GM plants showed a severe, abnormal phenotype (Hacham et al. 2002). In summary, in most studies, increasing flux through the methionine
biosynthetic pathway seems to have produced little increase in the methionine content of endogenous plant protein.

Photosynthetic source leaves are assumed to be the major sites of sulphur assimilation in plants; however, it has been demonstrated that the pathway of reductive sulphur assimilation is active in developing soya bean seeds (Sexton & Shibles 1999) and that sulphur amino acid biosynthesis occurs in developing embryos in the grain legume, *Lupinus angustifolius* (Tabe & Droux 2001). Thus sulphur assimilation in the developing seed itself appears to be an important source of sulphur amino acids for legume seed storage protein synthesis. Recently, manipulation of the cysteine biosynthetic pathway in developing lupin seeds was shown to result in large increases in free cysteine, although free methionine and total sulphur amino acid levels were not increased (L. Tabe, unpublished data).

**(b) Expression of methionine-rich proteins in GM plants**

Expression of an added gene for a methionine-rich protein or ‘methionine sink’ has been a successful GM strategy for modifying plant methionine content. This approach has been mainly used to improve the amino acid balance of legume seed protein, which can contain less than half the methionine required for optimal animal nutrition. Early attempts to increase the content of methionine in seeds by transgenic expression of genes for endogenous storage proteins mutated to add extra methionine residues were unsuccessful (e.g. Hoffman et al. 1988). A better strategy was the creation of a synthetic gene encoding an artificial protein rich in essential amino acids. Expression of a synthetic protein containing 31% lysine and 20% methionine residues in GM tobacco seeds under the control of a seed-specific promoter increased the total methionine concentration by 30% in the mature seeds (Keeler et al. 1997). A comparable result in a grain legume would give significant improvement in the nutritive value of the seed protein.

Methionine sink manipulation has most commonly involved transgenic expression of naturally occurring, methionine-rich plant proteins. Sulphur-rich proteins that have been expressed in GM dicots include 2S seed albumins from sunflower, Brazil nut and sesame, proteins that contain up to 18% methionine residues (Altenbach et al. 1989; Kortt et al. 1991; Tai et al. 1999a,b). This strategy has mainly been applied to the grain legumes owing to their low-intrinsic methionine concentrations; however, seeds of other species such as maize and canola have also been modified, not because they lack methionine themselves, but as a means of providing additional protein methionine in animal feed formulations containing grain legumes. For example, sulphur-rich zeins containing up to 28% methionine residues have been overexpressed in GM maize (Chui & Falco 1995).

The sunflower 2S seed albumin was used in a strategy to improve the sulphur amino acid content of seed protein in narrow leaf lupin (*Lupinus angustifolius*), a grain legume widely grown in Australia and used mainly for animal feed. The sunflower albumin was expressed under the control of a strong, seed-specific promoter from a pea vicilin gene in the GM lupins and resulted in increases of up to 100% in total seed methionine when compared with the parental genotype. The additional methionine was demonstrated to be available to rats and chickens (Molvig et al. 1997; Ravindran et al. 2002). Importantly, the methionine was also of benefit to sheep due to the rumen stability of the added methionine-rich sink protein (White et al. 2001). The Brazil nut 2S albumin has been expressed in a number of seeds including tobacco, canola, narbon bean and soya bean, with increases in total seed methionine of 30–100% when compared with wild-type (Altenbach et al. 1989, 1992; Muntz et al. 1997; Tabe & Higgins 1998). The levels of seed methionine in the GM soya beans and narbon beans were predicted to be sufficient for optimal animal nutrition; however, the potential human allergenicity of the Brazil nut protein has prevented it from being used commercially.

Expression of high-methionine proteins in GM cereals has met with mixed success. A sulphur-rich 2S albumin from sesame was reported to increase the total seed methionine by up to 75% in GM rice (Lee et al. 2003). In contrast, expression of the sunflower 2S albumin in GM rice produced no significant increase in seed methionine. In the latter case, endogenous seed protein composition changed in a way that resembled the well-characterized responses of seed proteins to plant sulphur nutritional stress (Hagan et al. 2003). In the GM rice grain expressing the sunflower protein, endogenous, sulphur-poor proteins were upregulated, while sulphur-rich proteins were downregulated. This apparent reallocation of limited sulphur reserves within the developing rice grain resulted in mature GM grain with different protein composition, but much the same sulphur amino acid concentration as the parental genotype. It is not clear why the expression of two very similar 2S albumins in rice, under the control of similar seed-specific promoters, should produce such contrasting outcomes. There are, however, a number of reports of compensatory changes in endogenous pools of sulphur in GM seeds expressing added, sulphur-rich proteins.

Individual kernels of GM maize overexpressing a high-methionine 10 kDa zein showed reduced levels of a separate endogenous sulphur-rich 12 kDa zein (Anthony et al. 1997). Likewise, endogenous sulphur-rich proteins were under-represented in GM soya bean seeds that accumulated the Brazil nut 2S protein (Jung et al. 1997). GM lupins expressing the sunflower albumin had reduced levels of transcripts encoding endogenous sulphur-rich seed storage proteins (Tabe & Droux 2002). The GM lupins also contained less oxidized sulphur than parental seeds grown in matched conditions. Similarly, GM narbon beans expressing the Brazil nut albumin contained smaller endogenous pools of sulphur in the form of the tri-peptide γ-glutamyl-S-ethenyl-cysteine than parental control seeds (Muntz et al. 1997). Thus, both protein and non-protein pools of sulphur were apparently deployed to supply methionine for the synthesis of the added sulphur sink protein in the GM seeds. In summary, it has certainly been possible to increase total seed methionine by plant genetic modification, although the evidence indicates that in many cases this has involved reallocation of...
endogenous pools of sulphur rather than increased delivery of sulphur to the seeds. In some cases, the data suggest that methionine enrichment has been achieved via increased rates of methionine biosynthesis in the developing seeds (Tabe & Droux 2002).

(c) Combined approaches

Manipulation of methionine biosynthesis in plants has greatly furthered the understanding of the regulation of flux through the pathway but, as a means of improving methionine content, this strategy suffers from the lack of stable storage of the additional methionine. On the other hand, addition of genes for methionine-rich storage proteins has produced GM seeds that in some cases are predicted to contain enough sulphur amino acids to satisfy the growth requirements of animals and humans. However, in other cases, the results indicate that methionine biosynthesis in developing seeds became limiting; for example in lupins, whose starting concentration of methionine was very low (Tabe & Higgins 1998; Tabe & Droux 2002). The obvious solution of combining the addition of a sulphur sink with modification of the sulphur amino acid biosynthetic pathway is the subject of current work. Some success has been reported; for example, expression of both the Brazil nut 2S albumin and a feedback-insensitive aspartate kinase gave additive increases in total methionine in seeds of GM narnon beans, although most of the effect was apparently due to the Brazil nut protein (Demidov et al. 2003). More recently, it has been reported that co-expression of an Arabidopsis CGS enzyme with a sulphur-rich zein in GM alfalfa leaves increased accumulation of the zein when compared with its expression alone in GM alfalfa (Bagga et al. 2005).

(d) High-methionine mutants

A number of plant mutants with increased levels of methionine have been isolated by selection on ethionine, a toxic analogue of methionine. Using this approach, three distinct groups of mutated genes have been characterized in A. thaliana, and have been found to define three enzymes from the methionine and $\beta$-adenosylmethionine biosynthetic pathways (Inaba et al. 1994; Chiba et al. 1999; Bartlem et al. 2000; Goto et al. 2002; Shen et al. 2002). A soya bean mutant with increased total methionine in its mature seeds was recently isolated using an initial screen for ethionine resistance. The outcome of this work was a soya bean variant that was predicted to supply enough methionine for optimal animal nutrition without supplementation with synthetic amino acid (Imsande 2001).

A natural maize mutant was identified by screening for germination on media containing lysine plus threonine, a combination that inhibits flux through the aspartate amino acid biosynthetic pathway, leading to methionine starvation. The mutant seeds had high levels of a specific, methionine-rich seed storage protein, the sulphur-rich $\delta$-zein. Analysis of the mutant revealed a lesion in a post-transcriptional control mechanism that normally suppressed $\delta$-zein transcript levels (Phillips & McClure 1985; Swarup et al. 1995). The same high-methionine phenotype was subsequently engineered in GM maize by mutation of the $\delta$-zein gene to remove the target site for negative regulation by the $\text{dze}1$ locus. The modified maize had methionine levels theoretically high enough to obviate the need for synthetic methionine in animal feed formulations containing the GM seed (Lai & Messing 2002).

(e) Prospects

In summary, both mutation breeding and genetic modifications have been used successfully to improve the content of the nutritionally essential sulphur-containing amino acid methionine in plants. In both cases, modified plant products with changed seed storage protein composition would be screened for changes in allergenicity before commercial release, since many seed proteins elicit allergic responses in some people (Mills et al. 2003). The goal of increasing methionine content, and hence nutritive value, of plant protein is currently being achieved and will no doubt continue to progress in the near future.

4. STARCH BIOSYNTHESIS AND FUNCTIONALITY

The synthesis of starch has fascinated researchers for several decades owing to the paradox between the apparent structural simplicity of starch, yet its synthetic complexity. The apparent simplicity of its structure arises because starch is composed of a single monomer, glucose, linked together into polymers through just two linkage types, $\alpha$-1,4 and $\alpha$-1,6. However, the heterogeneity of chain lengths and total molecular weight distribution, plus heterogeneity in the number and placement of $\alpha$-1,6 linkages leads to starches being composed of polydispersse populations of molecules, with each population having different functional properties. Adding further to the complexity, starches are laid down in granules, and the control of granule size, number and structure adds a further layer through which functional properties are determined. The populations of molecules within a given starch can be classified into two groups: amylose, a relatively linear $\alpha$-1,4 glucan of total degree of polymerization from 500 to 2000 and fewer than 1% $\alpha$-1,6 branch points, and amyllopectin, a highly branched molecule (3–4% $\alpha$-1,6 linkages) with a high molecular weight (degree of polymerization 5000—50 000).

The initial committed step in starch synthesis is the formation of ADP-glucose from glucose-1-phosphate and ATP. This step is unique to starch synthesis compared with other metabolic needs. In the cereal grain, it has long been recognized that the enzyme catalysing this step, ADP-glucose pyrophosphorylase, is composed of two types of subunits, ‘large’ and ‘small’ (Morell et al. 1987). The activity of this enzyme is regulated at three distinct levels. Firstly, the enzyme is present in both cytosolic and plastidic forms (Deyner et al. 1996; Thorbjornsen et al. 1996). In developing endosperm the majority of the flux is via the cytosolic form, while in chloroplasts the plastidic form dominates. Secondly, the enzyme is subject to redox control, apparently coordinating activity levels with photosynthetic flux. Thirdly, the enzyme is subject to complex allosteric regulation, being activated by 3-phosphoglycerate and inhibited by inorganic

Phil. Trans. R. Soc. B (2008)
phosphate (Ghosh & Preiss 1966). Understanding how these regulatory mechanisms interact to modulate the flux through the starch synthesis pathway remains an area of ongoing study.

The synthesis of amylose requires the activity of granule-bound starch synthase (GBSS), an enzyme that is principally located within the starch granule. There is evidence that other enzymes contribute to the synthesis of amylose, however GBSS is the only enzyme absolutely required for its synthesis (Ball & Morell 2003). There are separate GBSS genes expressed in endosperm and other parts of the plant providing a basis for observed differences in amylose content and structure between leaf and endosperm starches (Nakamura et al. 1998; Edwards et al. 2002).

The synthesis of amylopectin is complex, with a range of enzymes contributing. Firstly, plants contain a family of starch synthases with differing substrate specificities responsible for the elongation of amylopectin chains. Genetic analysis suggests that these isoforms have differing roles in amylopectin synthesis. Starch synthase (SS) I is thought to be responsible for the synthesis of the short external chains of amylopectin (DP6–10; Delvalle et al. 2005), whereas SSIIa is responsible for the synthesis of longer chains, from DP12–20. Elimination of this enzyme in barley (Morell et al. 2003), wheat (Yamamori et al. 2000) and rice (Umemoto et al. 2002) leads to a very characteristic phenotype involving reduced amylopectin external chain length, reduced granule gelatinization temperature and reduced starch swelling properties. The role of SSIII is less clear but this enzyme, along with GBSS, contributes to the synthesis of longer chains present in amylopectin (Gao et al. 1998; Zhang et al. 2005). There are at least two other classes of starch synthase genes present in the rice genome, SSIIb and SSIV. Both are primarily expressed in leaves and their roles are currently being defined.

In monocot plants, three branching enzyme genes are found, branching enzyme (BE) I, BEIIa and BEIIb. Mutation studies in a range of species indicate that the effects of eliminating BEI activity in a normal background range from undetectable to extremely subtle (Blauth et al. 2002; Satoh et al. 2003; Regina et al. 2004). Effects of BEI mutations are only seen in a background lacking either BEIIa or BEIIb. In maize, mutants in each of the genes have been identified and double mutants constructed. Mutation of the BEIIa gene shows that there is no detectable effect on amylose content or starch structure in the endosperm, but there is a dramatic effect on leaf starch. Mutations in BEIIb have long been known to result in a high-amylose phenotype, in keeping with the observation that this is a dramatic effect on leaf starch. Mutations in BEIIb are responsible for the synthesis of amylose, however GBSS is the only enzyme absolutely required for its synthesis (Ball & Morell 2003). There are separate GBSS genes expressed in endosperm and other parts of the plant providing a basis for observed differences in amylose content and structure between leaf and endosperm starches (Nakamura et al. 1998; Edwards et al. 2002).

The synthesis of amylopectin is complex, with a range of enzymes contributing. Firstly, plants contain a family of starch synthases with differing substrate specificities responsible for the elongation of amylopectin chains. Genetic analysis suggests that these isoforms have differing roles in amylopectin synthesis. Starch synthase (SS) I is thought to be responsible for the synthesis of the short external chains of amylopectin (DP6–10; Delvalle et al. 2005), whereas SSIIa is responsible for the synthesis of longer chains, from DP12–20. Elimination of this enzyme in barley (Morell et al. 2003), wheat (Yamamori et al. 2000) and rice (Umemoto et al. 2002) leads to a very characteristic phenotype involving reduced amylopectin external chain length, reduced granule gelatinization temperature and reduced starch swelling properties. The role of SSIII is less clear but this enzyme, along with GBSS, contributes to the synthesis of longer chains present in amylopectin (Gao et al. 1998; Zhang et al. 2005). There are at least two other classes of starch synthase genes present in the rice genome, SSIIb and SSIV. Both are primarily expressed in leaves and their roles are currently being defined.

In monocot plants, three branching enzyme genes are found, branching enzyme (BE) I, BEIIa and BEIIb. Mutation studies in a range of species indicate that the effects of eliminating BEI activity in a normal background range from undetectable to extremely subtle (Blauth et al. 2002; Satoh et al. 2003; Regina et al. 2004). Effects of BEI mutations are only seen in a background lacking either BEIIa or BEIIb. In maize, mutants in each of the genes have been identified and double mutants constructed. Mutation of the BEIIa gene shows that there is no detectable effect on amylose content or starch structure in the endosperm, but there is a dramatic effect on leaf starch. Mutations in BEIIb have long been known to result in a high-amylose phenotype, in keeping with the observation that this is a dramatic effect on leaf starch. Mutations in BEIIb are responsible for the synthesis of amylose, however GBSS is the only enzyme absolutely required for its synthesis (Ball & Morell 2003). There are separate GBSS genes expressed in endosperm and other parts of the plant providing a basis for observed differences in amylose content and structure between leaf and endosperm starches (Nakamura et al. 1998; Edwards et al. 2002).

The synthesis of amylopectin is complex, with a range of enzymes contributing. Firstly, plants contain a family of starch synthases with differing substrate specificities responsible for the elongation of amylopectin chains. Genetic analysis suggests that these isoforms have differing roles in amylopectin synthesis. Starch synthase (SS) I is thought to be responsible for the synthesis of the short external chains of amylopectin (DP6–10; Delvalle et al. 2005), whereas SSIIa is responsible for the synthesis of longer chains, from DP12–20. Elimination of this enzyme in barley (Morell et al. 2003), wheat (Yamamori et al. 2000) and rice (Umemoto et al. 2002) leads to a very characteristic phenotype involving reduced amylopectin external chain length, reduced granule gelatinization temperature and reduced starch swelling properties. The role of SSIII is less clear but this enzyme, along with GBSS, contributes to the synthesis of longer chains present in amylopectin (Gao et al. 1998; Zhang et al. 2005). There are at least two other classes of starch synthase genes present in the rice genome, SSIIb and SSIV. Both are primarily expressed in leaves and their roles are currently being defined.

In monocot plants, three branching enzyme genes are found, branching enzyme (BE) I, BEIIa and BEIIb. Mutation studies in a range of species indicate that the effects of eliminating BEI activity in a normal background range from undetectable to extremely subtle (Blauth et al. 2002; Satoh et al. 2003; Regina et al. 2004). Effects of BEI mutations are only seen in a background lacking either BEIIa or BEIIb. In maize, mutants in each of the genes have been identified and double mutants constructed. Mutation of the BEIIa gene shows that there is no detectable effect on amylose content or starch structure in the endosperm, but there is a dramatic effect on leaf starch. Mutations in BEIIb have long been known to result in a high-amylose phenotype, in keeping with the observation that this is a major BEII isoform expressed in the endosperm. Recently, Regina et al. (2006) have demonstrated that in wheat (and barley, unpublished data), BEIIa is more highly expressed than BEIIb and suppression of BEIIa, rather than BEIIb, is critical to achieve increased amylose content.

A conundrum in starch synthesis research is the role of debranching enzymes. Genome sequence studies in a wide range of plants show that there are four debranching enzyme genes in the plant genome, three isoamylase-like genes (isoamylases 1, 2 and 3) and one pullulanase- (or limit dextrinase-)type gene (Morell & Myers 2005). Mutation studies in a range of species, including rice (Nakamura et al. 1996), maize (James et al. 1995), barley (Burton et al. 2002), Arabidopsis (Zeeman et al. 1998) and Chlamydomonas (Mouille et al. 1996), demonstrate that mutation in isoamylase 1 leads to a low-starch high-phytoglycogen phenotype. More recent data suggest that an identical phenotype is recovered when isoamylase 2 is mutated and it is suggested that this is because isoamylase 1 and 2 form a complex whose function is abolished if either is absent. The role of isoamylase 3 remains unclear. Pullulanase mutants have only a subtle direct phenotype (Dinges et al. 2003) but have major effects in an isoamylase 1-deficient background, indicating that there may be some functional overlap between the two debranching enzymes. The role played by these various debranching enzymes in starch biosynthesis remains a matter of debate. One view is that isoamylases are directly involved in starch synthesis, 'editing' the emerging amylopectin molecule such that a crystallization-competent amylopectin is formed in the crystalline lamellae regions of the starch granule (Myers et al. 2000). This view is supported by observations that relate the level of activity of isoamylase in the developing endosperm to corresponding changes in branch point frequency and starch structure in starch granules (Kubo et al. 2004). Other views are that isoamylase plays a role in removing highly branched phytoglycogen from the amyloplast stroma (Zeeman et al. 1998) and debranching enzymes are involved in starch granule initiation (Burton et al. 2002).

Despite this wealth of information on the starch synthesis pathway, there are still glaring gaps in our knowledge. The synthesis of bacterial glycogen involves the same core enzyme activities (ADP–glucose pyrophosphorylase, glycogen synthase, branching enzyme and a debranching enzyme) as higher plant starch synthesis, yet a very different non-crystalline product is synthesized by bacteria (Ball & Morell 2003; Dauvillee et al. 2005). Studies of diverse green algae show that a complex set of starch synthesis isoforms is present in even the simplest green algae, indicating the high conservation of function of the various isoforms (Rai et al. 2004). Interestingly, Nakamura et al. (2005) have identified cyanobacteria with semi-crystalline amylopectin, but a reduction in isoform number. The precise roles of individual isoforms, and their interactions, remain to be dissected.

There is still a paucity of direct information on the events that lead to starch granule initiation, and little understanding of how the complex granule developmental processes seen in wheat and barley starches are controlled. One area of recent work that may provide a key to unlock further secrets in starch biosynthesis is research describing the presence of phosphorylation-dependent complexes of starch biosynthetic enzymes in developing cereal endosperm (Tetlow et al. 2004a). Complexes between starch biosynthetic enzymes have the potential to channel substrates to specific structural endpoints acting as 'carbohydrate chaperones' (Tetlow et al. 2004b). Further research is required to determine how the various levels of regulation, transcriptional, allosteric and post-translational, intersect to control the fine structure of starch and the structure of starch.
granules. Only when this level of knowledge is achieved, the full potential for the rational design of starches with specific functionality will be possible (Morell & Myers 2005).

5. MANIPULATING SEED FATTY ACIDS FOR HUMAN NUTRITION AND FOR INDUSTRY

Crop and livestock production systems are the mainstay of the many essential nutrients that support human life, health and well-being. As more is being learnt about the specific role of key nutrients in human nutrition, it is also becoming apparent that the supply of some nutrients is compromised and in some cases may not be sustainable into the future from current resources. The most notable of these potential shortfalls relate to the long chain polyunsaturated fatty acids (LC-PUFA) of the omega-3 (ω3) class, such as eicosapentaenoic acid (EPA, 20 : 5Δ4,8,11,14,17) and docosahexaenoic acid (DHA, 22 : 6Δ4,7,10,13,16,19), that are found predominantly in fish and other seafood. Inadequate levels of EPA and DHA are typical in Western-style diets that are low in seafood and have been associated with increased incidence of cardiovascular disease, cancer, stroke, diabetes, inflammatory disease, neuropsychiatric disorders and many other conditions prevalent in Western societies (Simopoulos 2003). Consequently, nutritionists and health authorities now regularly recommend significant increases in consumption of fish and other seafood rich in EPA and DHA. However, it is now widely acknowledged that global fisheries are fully exploited, with many on the verge of collapse (Myers & Worm 2003), and they may be inadequate to sustain even current levels of fish consumption. Fish farming and other forms of aquaculture are rapidly expanding and can help to overcome the declining catch from wild fisheries, but many aquaculture systems rely heavily on wild fisheries for feeds and are actually net consumers, not producers, of ω3 LC-PUFA (Naylor et al. 2000; Pauly et al. 2002). This situation means that existing marine-based sources of ω3 LC-PUFA are unlikely to be sufficient to sustain current levels and anticipated future increases in human needs.

Fortuitously, the advent of genetic engineering technologies is now providing a solution to this dilemma through the development of transgenic plants equipped with the ability to synthesize ω3 LC-PUFA. This is being achieved by the transfer of genes encoding the EPA and DHA biosynthetic pathways from marine microalgae and other micro-organisms into agricultural crops, in particular oilseed crops. All higher plants have the ability to synthesize the main C18-PUFA, linoleic acid (LA, 18 : 2Δ9,12) and ω3-linolenic acid (ALA, 18 : 3Δ9,12,15), and some can also synthesize ω6-linolenic acid (GLA, 18 : 3Δ6,9,12) and stearidonic acid (SDA, 18 : 4Δ6,9,12,15). However, higher plants are unable to further elongate and desaturate these ω3 C18-PUFA to produce ω3 LC-PUFA that are characteristic of the marine microalgae that are the ultimate source of EPA and DHA found in fishes. Synthesis of ω3 LC-PUFA in higher plants therefore requires the introduction of genes encoding all of the biosynthetic enzymes required to convert ALA into EPA and DHA. Substantial parallel gene discovery efforts conducted over the last 10 years in a range of LC-PUFA-synthesizing organisms have resulted in the cloning of genes for all of the fatty acid desaturase and elongase enzymes involved in the aerobic pathway for LC-PUFA synthesis and have been reviewed in detail (Sayanova & Napier 2004). Recently, significant progress has been reported in expressing these pathways transgenically in seeds with the achievement of substantial levels of EPA (20% of total fatty acids) in soya bean seed oil (Kinney et al. 2004) and later the synthesis of low levels of DHA (1–2% of total fatty acids) in A. thaliana (Robert et al. 2005) and Brassica juncea (Wu et al. 2005). These studies used different combinations of LC-PUFA biosynthetic genes from a variety of organisms and revealed the considerable complexity associated with introduction of this multi-step fatty acid biosynthetic pathway into higher plants (Singh et al. 2005). It is probable that additional or alternative metabolic manipulations will be required in order to achieve significantly higher levels of DHA synthesis and accumulation in transgenic seed oils. However, it is now clearly apparent that seeds can be engineered to produce the range of ω3 LC-PUFA required in the human diet and potentially in concentrations that should be nutritionally effective. Crop plants engineered in this way will ultimately provide the affordable, renewable and sustainable sources of ω3 LC-PUFA that are urgently needed to overcome the inadequate and potentially unsustainable supply from traditional marine sources.

(a) Sustainable industrial raw materials supply

As well as providing the capability to achieve a sustainable increase in the supply of nutritional oils, genetic manipulation of fatty acid metabolic pathways in plants can also open the way for a more sustainable supply of industrial raw materials, by enabling these to be sourced from renewable plant resources rather than from increasingly scarce and non-renewable petroleum. The recent persistent escalation in the price of petroleum and predominantly pessimistic supply forecasts have driven a considerable expansion in the use of plant-based fuels, such as ethanol and biodiesel, as commodity scale alternatives to conventional fuels. It is anticipated that in the future other high-value speciality industrial products currently produced by the petrochemical industry will be produced on a renewable basis from oleochemical sources, predominantly from plants producing specific molecular structures required as starting materials for advanced chemicals and polymers. These products will be generated by metabolic engineering of plant biosynthetic pathways either by redirecting pathways towards the accumulation of current intermediate compounds, such as in the production of lauric acid (C12 : 0) in rapeseed (Voelker et al. 1996), or by the introduction of new biosynthetic pathways that lead to completely novel end products, such as the production of polyhydroxyalkanoates in various plant tissues (Poirier 1999). In this regard, the engineering of fatty acid metabolic pathways in oilseeds is likely to be a particularly fruitful area, due to the similarity of acyl chains to petrochemically derived hydrocarbons and
their ability to be functionally derivatized by a wide array of acyl-modifying enzymes.

Because they have been selected and bred mainly for food purposes, our major oilseeds are very restricted in the range of fatty acids that they contain, usually only five (palmitic, stearic, oleic, linoleic and linolenic). However, in nature, there is an enormous diversity of fatty acid structures (Badami & Patil 1981), including many functionalities such as hydroxylation, epoxidation, acetylenation and conjugation, that impart properties required for specific industrial uses. Gene technology has enabled the enzymes responsible for these functionalities to be cloned from various sources and expressed transgenically in oil-accumulating crop species in order to develop novel industrial oils.

To date most attention has been focused on C18 fatty acids that are modified at the Δ12 position by the addition of epoxy or hydroxy groups, or by the formation of triple bonds (acytlenic) or conjugated double bonds. The introduction of these functionalities into C18 fatty acids are catalysed by a family of divergent forms of the fatty acid Δ12-desaturase (FAD2) enzyme. FAD2-like genes encoding Δ12 epoxygenases, hydroxylases, acetylenases and conjugases have all been cloned (several years ago) and recently reviewed (Jaworski & Cahoon 2003). Transgenic expression of these divergent FAD2 genes in Arabidopsis and other oil-accumulating seeds has generally resulted in synthesis of the Δ12-modified fatty acid, but in disappointingly low concentrations (less than 10% of oil), even though the modified fatty acids are present at very high concentrations in the source plants (60–90%). For example, vernolic acid, a Δ12-epoxygenated C18 fatty acid present in several wild plant species, has been produced in transgenic plants by expression of the fatty acid Δ12-epoxygenase enzyme obtained for a number of plant sources including Crepis palaestina (Singh et al. 2000a,b), Euphorbia lagascae (Cahoon et al. 2002) and Stokesia laevis (Hatanaka et al. 2004). However, in each case, the level of vernolic acid synthesis was initially low regardless of whether the Δ12-epoxygenase was a divergent FAD2 type such as from Crepis palaestina or a cytochrome P450 type such as from Euphorbia lagascae. It has subsequently been demonstrated that the level of vernolic acid synthesized in Arabidopsis seeds expressing the Crepis palaestina FAD2-like Δ12-epoxygenase can be enhanced from initial levels of approximately 6% (Singh et al. 2000a,b) to approximately 20% of total fatty acids (Zhou et al. 2006) by increasing the availability of linoleic acid substrate. This was achieved by co-expressing the Δ12-epoxygenase with additional Δ12-desaturase genes in a mutant Arabidopsis genotype lacking the fatty acid elongase (FAE1) and Δ15-desaturase (FAE3) enzymes that would otherwise compete for substrates involved in synthesis of Δ12-epoxy fatty acids.

These enhanced levels still fall well short of the high concentrations needed for industrial use, and it remains to be determined what additional manipulations may lead to high-level synthesis and accumulation of introduced Δ12-modified fatty acids. The common experiences of expressing Δ12-modifying enzymes in transgenic seeds reveal that plants vary considerably in the ability of their background metabolic machinery to handle the newly synthesized fatty acids. The novel fatty acids must be efficiently moved from their site of synthesis on PC and deposited in TAG to enable high-level accumulation in seed storage oils and exclusion from functional membrane lipids. Organisms that naturally accumulate these unusual fatty acids in abundance will have evolved appropriate metabolic pathways and substrate specificities to achieve these transfers efficiently. Increasing attention will, no doubt, therefore be placed on understanding the enzymatic steps and substrate specificities that such organisms use to achieve high-level synthesis and accumulation of these fatty acids, and on cloning genes for the enzymes involved. It is probable that this will uncover genes for specialized forms of the various acyltransferase and TAG assembly enzymes capable of efficiently handling the unusual fatty acids. Co-expression of such genes along with the previously introduced fatty acid biosynthetic pathways should contribute to further increases in accumulation of novel fatty acids in transgenic plants in the future and lead to the development of economically viable crop sources of industrial raw materials.

6. DISCOVERY AND USE OF GENES FOR IMPROVED DISEASE RESISTANCE IN CROP PLANTS

The use of disease-resistant crop cultivars provides an effective method of controlling a large number of diseases. However, continuous breeding efforts are required to counter evolution or migration of new pathogen strains. One stumbling block continues to be the lack of agreement regionally between breeders as to the most effective deployment of valuable R genes to prevent their stepwise erosion by pathogen evolution. Plant molecular biology is and will make increasing contributions to resistance breeding by making resistance breeding more effective and more efficient, especially through the use of markers for breeding and providing resistance genotypes for varieties to improve decision making about their deployment.

(a) DNA markers for breeding

Our efforts have been mainly targeted at rust, nematode diseases of cereals and barley yellow dwarf virus, and molecular markers have been developed for improved breeding efficiency. Effective genetic resistance in wheat for cereal cyst nematodes is currently provided by the Cre1 and Cre3 genes. Breeding new resistant varieties has, however, been hindered by the slow and laborious nature of the plant bioassay for nematode resistance. DNA markers have now been identified for both resistance genes based on cloned genes of the nucleotide binding site–leucine rich repeat disease resistance gene class (deMajnik et al. 2003). These genes co-segregate with the Cre1 and Cre3 resistance genes and although there is no direct evidence to indicate that the cloned genes themselves control nematode resistance, they have provided excellent sources for development of simple, rapid and accurate PCR-based markers that are currently being used by wheat breeders.
Wheat breeding has relied heavily on genetic resistance to rust disease to control stem, stripe and leaf rust. Breeding efforts have been particularly successful for stem rust using major genes for resistance and DNA markers for resistance are being increasingly used. In areas where stem rust resistance has been a major breeding objective, success has been achieved mainly by using varieties carrying several different stem rust resistance genes, diversity of resistance genotypes and discouragement of the cultivation of susceptible varieties. DNA markers are now being used increasingly for these breeding efforts. DNA markers need to be simple to use and also applicable to as wide a range of breeders germplasm as possible. For example, while some markers can be useful for genetic mapping of resistance genes in particular crosses, they are frequently not useful in all breeder lines where they fail to detect polymorphisms between resistance gene donors and susceptible recurrent parents. Consequently, there can be a long development stage between marker identification and application that involves fine-tuning to produce a robust marker across a range of useful genotypes.

Many wheat varieties carry the durable stem rust resistance gene Sr2 that is effective in providing partial resistance against all strains of stem rust at the adult stage of growth. PCR-based DNA markers have now been developed for marker-assisted breeding using the Sr2 gene (Spielmeyer et al., 2003), and have provided an entry point to finely map this gene for future molecular cloning (Kota et al., 2006) with the aim of understanding the molecular basis of an adult plant, durable, non-strain-specific resistance gene. Several other stem rust resistance gene markers have been developed and are described below. Good progress is being made in developing a PCR-based marker for the durable adult plant leaf and stripe rust gene pair Lr34–Yr18.

(b) DNA Markers useful for gene stacking
Pyramids or gene stacks of multiple stem rust resistance genes in a single variety can provide durable resistance. Traditionally, R gene pyramids are achieved using sequential bioassays with rust strains capable of differentiating those different resistance genes. This becomes more difficult for breeders if each of the genes used provide resistance to all available pathogen strains. This is where DNA markers will make a big contribution to providing simple tests for the presence of specific R genes. For stem rust, markers for Sr38, Sr24, Sr26, SrR and Sr31 have now been developed (Seah et al., 2001; Mago et al., 2005a,b). The latter four genes provide resistance to all stem rust strains currently found in Australia and the markers for Sr24 and Sr26 that provide resistance to the proliferating strain Ug99 now found in Africa will have global applications.

(c) DNA Markers for ‘value adding’ to alien resistance sources
Many of the currently effective stem rust resistance genes are derived from wheat relatives and many have negative dough characteristics that are physically linked to the same chromosome region as the resistance genes. They are consequently not suitable for use in high-quality bread wheats. For several of these R gene sources, the flanking alien chromatin regions have been reduced by recombination in ph16 mutant background (Łukaszewski, 2000). DNA markers are also being used to detect recombinants carrying the R gene, but with reduced alien flanking DNA (Rogowsky et al., 1991). Retained DNA markers are being used for the deployment of the modified sources of Sr31, SrR and Sr26 to produce near-isogenic lines for assessment of yield and quality effects and introduction as pyramids into adapted cultivars.

(d) Cloned rust resistance genes
The first rust resistance genes have been cloned from flax (Lawrence et al., 1995) and more recently from cereals (Collins et al., 1999; Brueggeman et al., 2002; Feuillet et al., 2003; Huang et al., 2003). Apart from providing the first insights into how rust resistance genes function, cloned genes will make a positive impact on plant breeding.

An interesting and valuable rust resistance gene for stem rust Rpg1 has been cloned from barley (Brueggeman et al., 2002). This gene, which is not from the most common NBS—LRR class of plant disease resistance genes, has provided durable stem rust resistance in barley. While barley plants transgenic for this gene provide an even higher level of resistance than the natural sources of the gene (Horvath et al., 2003), no experiments reporting the function of this gene in wheat have yet been reported.

Initial observations with cloned disease resistance transgenes indicated that they might only function in species closely related to the source plant (Tai et al., 1999a,b). More recent data show this is not necessarily the case. When co-expressed in tobacco, the flax rust resistance protein L6 recognizes the corresponding flax rust avirulence protein AvrL567 and induces a hypersensitive response characteristic of a disease resistance reaction. This is likely to be due to direct interaction of the resistance protein and the avirulence protein (Dodds et al., 2004). Whether the gene functions in tobacco to give rust resistance is not possible to determine because tobacco is a non-host for the flax rust. Nevertheless, the transfer from the Linaceae family to the Solanaceae family shows that wide transfers of resistance genes between species can function.

When the current regulatory and political blockages to GM versions of food crops like wheat and barley are removed, a number of possibilities for GM resistance breeding should become available. For example, in barley and wheat, many specificities for powdery mildew occur at the Mla and PM3 resistance loci, respectively (Shen et al., 2003; Srichumpa et al., 2005). Cloning studies have shown that these are allelic and so cannot be easily recombined to produce gene pyramids for stable resistance—only one allele at a time can be deployed in a homozygous line. This nexus could be broken using transgenic plants and multiple R transgenes can be transferred to wheat or barley to make otherwise unobtainable resistance gene pyramids.

In our own work in stem rust resistance in wheat, cloned genes from cereals are providing perfect markers for breeding. Furthermore, we are aiming to clone three or more resistance genes, package them into a single gene construct and introduce them into wheat.
using Agrobacterium. Two advantages over traditional methods are envisaged. Firstly, using cloned genes, the effect of linked genes with quality and yield defects can be removed. Secondly, by packing them in a single transgene cassette, the three genes will segregate during breeding as a single unit. Using traditional breeding, individual progeny plants homozygous for three unrelated genes are rare in segregating families. So far, cloning R genes from large cereal genomes is still difficult, but technology is advancing rapidly with increasing genome sequence data available. Rust resistance breeding in cereals is set to make a big jump with both marker-assisted and transgenic breeding. Furthermore, biotechnology can deliver surprises and the recent reports that round-up ready wheat shows high levels of rust resistance after spraying with glyphosate provides a challenge to develop agronomic practices for wheat that combine both weed and rust control using round-up (Anderson & Kolmer 2005; Feng et al. 2005).

7. GM INSECT PROTECTED COTTON: AN AUSTRALIAN EXAMPLE OF TRANSGENIC PLANT IMPROVEMENT

By the mid 1990s, the Australian cotton industry was beginning to stretch at the seams as this relatively young agricultural enterprise began to experience difficulties in containing its main insect pests, two caterpillars of the Helicoverpa complex. Failures in pest control were not new to cotton and it was still fresh in the minds of many how the fledgling industry in the Ord Irrigation Scheme (in the far north of Australia) had gone into a catastrophic spiral of insecticide resistance and increasing pesticide application in the 1960s. This resulted in the use of 35 or more insecticide sprays per crop per season, still without reaching any profitable level of production. This unsustainable dependence on pesticides resulted in the closure of cotton production in that region in the 1970s and the transfer of Australia’s efforts to Eastern Australia where pest pressures were still high, but not as extreme as in the more tropical North. Despite its sensitivity to drought and the variable availability of irrigation water, which results in periodic reductions in overall output, cotton production in the East has worked well for many years, climbing to the edge of the GM plots to act as a decoy for foraging Helicoverpa armigera and Helicoverpa punctigera, the two main insects being controlled by 80% of the pesticides then applied to cotton. CSIRO played a central role in the breeding of the new insecticidal trait (Cry1Ac, sold under the Ingard brand name in Australia) into adapted, high-performing germplasm for Australia, its subsequent deployment and the research that underpinned the management strategies and agricultural practices needed to make it a sustainable pest management tool. At the time of its introduction, the industry was already undergoing some critical self-evaluation about its environmental practices and had instituted many reforms that were already having a impact on reducing pesticide usage, including the introduction of best management practice (BMP) into cotton production and appropriate certification of individual and corporate growers (CRDC 2003). By 2002, 60% of the Australian cotton crop was produced under BMP and incorporated the use of the GM insect protected varieties being developed by CSIRO with the Monsanto genes included in this genome.

The Ingard genes were introduced into Australia as cotton seed in the variety Coker 312 (an obsolete Texan variety, one of few cotton varieties amenable to genetic transformation and regeneration) that was itself unsuited for growth under Australian environmental and agricultural production conditions. Conventional backcross breeding was used to improve the germplasm base of the GM cotton by repeated backcrossing to elite CSIRO varieties that were among the best in the world for yield, fibre quality and disease tolerance. Multi-site evaluation across the cotton production area ensured that the new GM versions were well adapted and retained the high yield and other qualities of their recurrent parents. By 1996, CSIRO had produced sufficient seed of five Ingard varieties for an initial trial planting of approximately 40 000 ha. In the meantime, researchers were gathering all the necessary data for regulatory approval, crop agronomy and resistance management that were a necessary precursor to any commercial scale use of the new technology.

Regulation of GM products in Australia was handled by a two-component system that included an voluntary advisory panel of scientists (the Genetic Manipulation Advisory Committee) who assessed the safety of GM products and provided advice to a variety of State and Federal Statutory Agencies with responsibilities for particular areas of regulation of human health, food safety, occupational safety and the environment. Releases of GM cotton into the environment started in 1992 with a release of a few hundred plants within a large field of conventional cotton that served as a pollen trap and isolated the GM cotton from other cotton being grown in the area. Subsequent trials increased steadily in size to allow further pollen movement studies, efficacy assessments, breeding selections and seed increase, as well as the ecological impact studies required by regulators.

Pollen flow studies indicated that cotton was easily contained within trials (cotton being a predominantly in-breeding plant) and required a relatively modest surrounding buffer crop extending only 20 m beyond the edge of the GM plots to act as a decoy for foraging insects such as bees that were the most likely vectors of pollen dispersal (Llewellyn & Fitt 1996). Efficacy of pest control was not absolute and although it proved to

Phil. Trans. R. Soc. B (2008)

http://rstb.royalsocietypublishing.org/ on July 1, 2017
be high during the first part of the growing season, it was noted to decline after flowering (Fitt 2004). This was subsequently shown to translate into commercial production with most of the savings in pesticide applications occurring during the first half of the season, where H. punctigera was the main pest.

Ecological impact studies measured any non-target impacts on the myriad of insects and other invertebrates that frequent cotton crops. In addition, the possibility of movement of the transgene out of cultivated cotton into native Gossypium species with a resultant disruption of the fine balance of these species was required to be assessed.

Given the existing knowledge on the host range of the delta-endotoxin, it was expected that the GM cotton plants would not have a negative impact on other invertebrates and this was borne out by extensive surveys of insect abundance in relatively large (10 ha) plots in replicated trials over a couple of years (Fitt & Wilson 2002). Impacts of the Ingard cotton were restricted to reductions in numbers of Helicoverpa larvae and other lepidopteran species known to be sensitive to the Cry1Ac protein, with a secondary effect on some lepidopteran-specific wasp parasites that normally feed within Helicoverpa caterpillars. Other beneficial insects tended to be more abundant in the Ingard cotton crops and were certainly much more abundant than in cotton crops sprayed with the conventional spectrum of pesticides normally used to control Helicoverpa species. Detailed genetic studies concluded that the risks of outcrossing to Australian native G or C genomic species, Gossypium sturtianum L., of the transgenes present in the GM cotton (AD genome allotetraploids) were negligible (Brown et al. 1997), although some of the K genome species in the more remote parts of Northern Australia might require further examination, should a cotton industry ever be established there.

The only major remaining concern of both growers and regulators was whether the technology would last beyond a couple of seasons if the target insect species could develop resistance to the insecticidal protein expressed in the plants. Previous research had reported resistance to Cry proteins in the Indian meal moth (Plodia interpunctella) and the diamondback moth (Plutella xylostella). Akhurst et al. (2003) were able, under laboratory conditions, to select a strain of H. armigera that was resistant to the toxicity of Cry1Ac proteins, so it was clear that target pests could possibly develop resistance to the active ingredient of Ingard cotton. The cotton industry had for years grappled with the problem of chemical insecticide resistance and was reluctant to see Ingard technology wasted. They set up a Transgenic and Insect Management Strategy committee to oversee the deployment of this new technology and make recommendations to both growers and regulators on all aspects of resistance management in an effort to preserve the new GM technology. Australian growers voluntarily adopted a strict area restriction on the use of the single gene Ingard cotton that saw every farm plant a maximum of 30% by area of Ingard varieties until such time as a second generation product was available that contained two different insecticidal toxins that would be more robust in countering any resistance development in the crop pests. This restriction was put in place to ensure that any resistance genes selected in the insects in the transgenic crops would not be fixed in the population, but would always find mates emerging from the non-transgenic crop that carry sensitive alleles for susceptibility to the Cry1Ac toxin and hence continually dilute out the resistance, keeping resistance allele gene frequencies very low within the target insect populations (e.g. Roush 1997). These so-called ‘refugia strategies’ require the presence of non-transgenic crops in close proximity to the GM crop and have been adopted around the world in a variety of crops carrying GM insecticidal traits; they are an important component of management to delay resistance to insecticidal genes. Other management components included specified planting and harvesting windows, obligate crop destruction after harvest to prevent regrowth and cultivation to destroy overwintering pupae. These strategies have been successful and no field resistance selected in GM crops has been reported in any Helicoverpa species or other target lepidopteran insects (e.g. Tabashnik et al. 2005).

By 2002, CSIRO had produced 15 different GM cotton varieties (combinations of Ingard and the herbicide-resistant Roundup Ready cotton) and continually updated their variety suite to keep pace with developments in conventional cotton germplasm. Despite changes from year to year in variety adoption, the 30% cap on Ingard cotton remained for 6–7 years during which time growers maximized the environmental benefits from the reduced pesticide spraying required on Ingard and in general used the new cottons on their more sensitive environmental sites close to towns, rivers or other dwellings where pesticide drift was likely to be a problem.

In 2003, CSIRO released a new suite of GM varieties that contained the Cry1Ac and a second insecticidal gene, Cry2Ab (also developed by Monsanto), that were sold as Bollgard II cotton. Bollgard II went through the same regulatory assessment as Ingard cotton, under a new regulatory regime that replaced the previous voluntary system. In 2000, the Australian government had put in place legislation to regulate biotechnology through a newly created statutory authority the Office of the Gene Technology Regulator. This represented a somewhat radical departure from previous systems as its primary goal was to put GM regulation on as open and transparent a footing as anywhere in the world. The requirement for accreditation and the issuing of licenses for the conduct of all GM research as well as a capacity for significant legal and monetary penalties have been put in place to ensure a high level of compliance by both research organizations and biotech and seed companies (as well as opponents of GM who might be tempted to interfere with field trials). Australia has not seen the fierce opposition to GM crops characteristic of European countries and GM cotton in particular has had a relatively straightforward introduction into agriculture (primarily because there was a strong desire for the technologies on the part of farmers and very obvious environmental benefits). The same has not been true for GM canola despite its success in Northern
GM canola foundered at a State political level, even though it was given Federal regulatory approval. Bollgard II cotton has done extremely well in Australia and within 2 years of its introduction constituted over 90% of all the cotton planted in this country, the majority of it as Bollgard II/Roundup Ready varieties that allowed growers better insect and weed control. The greater efficacy in the control of Lepidopteran pests and the presence in the cotton of two different insecticidal toxins offering greater protection against the development of resistance in the target pests have seen the removal of the planting area restrictions and a reduction in the sizes of the required refuges. Initial indications are that Bollgard II has slashed pesticide usage for Lepidopteran control by more than 80%.

One of the key developments with this new insect control technology is that it has fostered a greater adoption of integrated pest management in cotton, which is leading to even further reductions in pesticide usage (Wilson et al. 2004). The success of GM cotton in Australia has highlighted the value of GM solutions to agricultural sustainability and bodes well for future agbiotech products. Success will depend on the right genetics (getting the products into the right genetic backgrounds), the right management (researching the appropriate management scenarios to ensure the delivery of the benefits promised by the technology) and the right communication (making sure that the community, both the agricultural community and the wider community, are aware of the benefits) for the commercialization of those products.

8. CONCLUSION

The topics discussed in this paper present a set of examples of the ways in which genetic modification to the biological software of our major food and fibre production plants will continue to enhance the yield and sustainability of agricultural systems. DNA technology is now routinely used in plant improvement programmes with DNA sequence markers enhancing both the speed and the power of selection schemes. Our rapidly increasing knowledge of the functioning of crop genomes has already provided enhanced performance in conventional breeding programmes and although transgenic crops have not been welcomed in all parts of the world, they have already gained significant approval levels as judged by their use in approximately 4% of the production area globally and that area has been increasing substantially in each of the last 7 years.

These transgenic crops, including a fibre crop, cotton and the food and feed crops maize, soya bean and canola, have all been accepted in the various countries of the world in which they are grown and have entered successfully into markets. This represents a significant growth incorporation of transgenic modifications into breeding systems.

The understanding of the molecular bases of plant processes that we have gained from the advances in genomics and our increasing knowledge of gene regulation are opening up a new generation of breeding advances, both through transgenic breeding and conventional breeding. One of the advantages in many crops is that once precise breeding objectives have been defined by research that has used all the power of the new technologies, then breeders are able to use new diagnostic tools to achieve the desired objectives through conventional breeding programmes. This is providing a bridging period of improvement in plant breeding while our societies move towards general acceptance of transgenic tools in plant improvement programmes for our food, feed and fibre crops.

The examples in the paper range over improved environmental responses and improved protection against pests and pathogens together with improved nutritional value of crop products. There are likely to be many other possibilities for tailoring our crop species in the future. For example, breeders in the past have been able to adjust the architecture of plants to fit agricultural systems; breeding tomatoes for a single mechanical harvesting procedure is a dramatic example of plant architecture modification to suit a modern agricultural practice. We can expect these modifications to be more extensive than we have seen so far. The modifications may deal with the type of inflorescence, phyllotaxis, the way in which leaves respond to light in spatial and temporal modes, and there is a lot to be gained in modification of root systems to suit particular soils and their water and nutrient availabilities.

We will also profit from modification of internal architecture, the anatomy of plant tissues; for example, the ratio of palisade and spongy mesophyll leaf cells and the geometry of tissues in the root system are areas in which we can expect telling alterations.

Some of the examples we have discussed in the paper have specifically referred to challenges in Australian agriculture systems, but the points emphasized have general applicability to cropping systems around the world. In the case of transgenic cotton in Australia, one of the most important features is that behind the successful introduction and acceptance of the transgenic crop was the coordinate and packaged introduction of the new genetic make-up of the crop along with the new and mandatory ways of agronomic management. These were seen to be of extreme importance in introducing the value of the new technology to farmers. Farmers realized that it would be a huge loss if we were to waste this new powerful technology in the way that we wasted many of the advantages of the new pesticides in the recent past.

A reasonable conclusion is that genetic modification of crops, which has been so powerful and so rewarding in terms of yield and management of many of the major production species over the past few decades, will hold enormous potential in all of the crop species we deal with. We have an increasing knowledge and power to modulate the development and functional operation of crop plants so as to provide optimal performance in our agricultural production system environments.

Agricultural performance rests on the interactions of genetics, management and the environment. We have not always fully coped with these interactions, and production levels in many parts of the world have been less reliable than we might have hoped for. In many cases, the health status of the natural resources in the production areas have suffered and there has been great
concern by society as to the damage agricultural systems sometimes inflict on surrounding non-agricultural environments. But although the environmental challenges have been increasing in recent years, and continue to increase as a result of climate change and other factors operating on production systems, we can be confident that the new genetics is providing an increased ability to adjust the biological software of our principal production species. We can expect, in a variety of production environments, to have the genetic modifications, coupled with appropriate management regimes, to result in an increased efficiency and sustainability of agri-business.

REFERENCES


Akhtar, R., James, W., Bird, L. J. & Beard, C. 2003 Resistance to the Cry1Ac delta-endotoxin of *Bacillus thuringiensis* in the cotton bollworm, *Helicoverpa armigera* (Lepidoptera: Noctuidae). *J. Econ. Entomol.* 96, 1290–1299.


Leustek, T. & Saito, K. 1999 Sulfate transport and assimilation in plants. Plant Physiol. 120, 637–644. (doi:10.1104/pp.120.3.637)


