THE PRODUCTION OF RECOMBINANT PHARMACEUTICAL PROTEINS IN PLANTS

Julian K-C. Ma, Pascal M. W. Drake* and Paul Christou‡*

Imagine a world in which any protein, either naturally occurring or designed by man, could be produced safely, inexpensively and in almost unlimited quantities using only simple nutrients, water and sunlight. This could one day become reality as we learn to harness the power of plants for the production of recombinant proteins on an agricultural scale. Molecular farming in plants has already proven to be a successful way of producing a range of technical proteins. The first plant-derived recombinant pharmaceutical proteins are now approaching commercial approval, and many more are expected to follow.

GENETIC MODIFICATION

Proteins are widely used in research, medicine and industry, but the extraction of proteins from their natural sources can be difficult and expensive. Also, the use of pharmaceutical proteins from natural sources can pose risks. For example, many people have contracted diseases from contaminated blood products or hormones. Other proteins, such as SINGLE-CHAIN FV FRAGMENTS (scFvs), are not found naturally. So, a simple and inexpensive system that allows the large-scale production of safe recombinant proteins would be highly desirable. Traditional production systems that use microbial fermentation, insect and mammalian cell cultures, and transgenic animals have drawbacks in terms of cost, scalability, product safety and authenticity[1–3] (TABLE 1). Recent studies have shown that MOLECULAR FARMING in plants has many practical, economic and safety advantages compared with more conventional systems, and so the use of plants for large-scale protein synthesis is gaining wider acceptance[4,5].

In this review, we discuss the technological basis of molecular farming in plants, with a focus on proteins that can be used for diagnostic, therapeutic and prophylactic applications. We provide a broad account of the types of pharmaceutical protein that can be produced on a commercial scale and examine the different expression systems that are being developed. We consider the advantages and limitations of each system, with a focus on the biochemical constraints that need to be addressed for the technology to reach its full potential.

Recombinant proteins expressed in plants

Plants have provided humans with useful molecules for many centuries, but only in the past 20 years has it become possible to use plants for the production of specific HETEROLOGOUS proteins. The first pharmaceutically relevant protein made in plants was human growth hormone, which was expressed in transgenic tobacco in 1986 (REF. 6). In this study, the hormone was expressed as a fusion with the *Agrobacterium* nopaline synthase enzyme. Since then, many other human proteins have been produced in an increasingly diverse range of crops. In 1989, the first antibody was expressed in tobacco[7], which showed that plants could assemble complex functional glycoproteins with several subunits. The structural authenticity of plant-derived recombinant proteins was confirmed in 1992, when plants were used for the first time to produce an experimental vaccine: the hepatitis B virus (HBV) surface antigen[8]. In a further report, the same group showed that the vaccine produced in tobacco plants induced the expected immune response after it had been injected into mice[9]. More recently, the range of

SINGLE-CHAIN FV FRAGMENTS (scFvs). Monoclonal antibody derivatives that comprise a single polypeptide in which the variable regions of the heavy and light immunoglobulin chains are joined together by a flexible linker. scFvs are advantageous because only one transgene is required, and the molecules themselves are small and lack the effector functions of normal antibodies; however, a disadvantage is that they are univalent, whereas serum antibodies are divalent.

**Department of Infectious Diseases, St. George's Hospital Medical School, Cranmer Terrace, London SW17 0RE, UK.*
‡Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Grafschaft, Auf dem Aberg 1, 57392 Schmallenberg, Germany.
Correspondence to J.K-C.M.
e-mail: jma@sghms.ac.uk
doi:10.1038/nrg1177

Table 1 | **Comparison of production systems for recombinant human pharmaceutical proteins**

System	Overall cost	Production timescale	Scale-up capacity	Product quality	Glycosylation	Contamination risks	Storage cost
Bacteria	Low	Short	High	Low	None	Endotoxins	Moderate
Yeast	Medium	Medium	High	Medium	Incorrect	Low risk	Moderate
Mammalian cell culture	High	Long	Very low	Very high	Correct	Viruses, prions and oncogenic DNA	Expensive
Transgenic animals	High	Very long	Low	Very high	Correct	Viruses, prions and oncogenic DNA	Expensive
Plant cell cultures	Medium	Medium	Medium	High	Minor differences	Low risk	Moderate
Transgenic plants	Very low	Long	Very high	High	Minor differences	Low risk	Inexpensive

MOLECULAR FARMING
The large-scale production of recombinant proteins in living cells or organisms; frequently applied to the use of crop plants or domestic animals as expression hosts because of the allusion to agriculture.

HETEROLOGOUS
In the context of this article, a gene or protein that is not derived from the species in which it is expressed.

TRANSPLASTOMIC
A transgenic plant in which the transgene is found in the plastid genome rather than the nuclear genome.

DIABODY
A recombinant antibody that comprises the heavy- and light-chain variable regions joined by a flexible peptide linker. The linker is long enough to allow separation of the domains so that two of the polypeptides can assemble into a dimer, making the antibody divalent.

MINIBODY
A recombinant antibody in which the heavy- and light-chain variable regions are part of the same polypeptide chain, which also includes the heavy-chain hinge region and one heavy-chain constant domain.

AGROINFILTRATED LEAVES
Usually leaves of tobacco (although many other species can be used) that are transiently transformed with *Agrobacterium tumefaciens*, which results in the transient expression of recombinant proteins. This is a useful strategy for testing expression constructs and obtaining small amounts of protein for analysis before going to the expense of transgenics.

recombinant proteins made in plants has extended to include industrial enzymes[10], technical proteins that are used in research[11], milk proteins that are suitable nutritional supplements[12] and new protein polymers with both medical and industrial uses[13]. Examples of pharmaceutical proteins that have been produced in plants are listed in TABLE 2.

Proteins as pharmaceuticals. Many pharmaceutical proteins of mammalian origin have been synthesized in plants. These range from blood products, such as human serum albumin for which there is an annual demand of more than 500 tonnes, to cytokines and other signalling molecules that are required in much smaller amounts. Until recently, most plant-derived proteins have been produced in transgenic tobacco and extracted directly from leaves. Generally, these proteins are produced at low levels, typically less than 0.1% of the total soluble protein. This low level of production probably reflects a combination of factors, with poor protein folding and stability among the most important.

In the past few years, the tobacco chloroplast system has been used to express human proteins at much higher levels. Human growth hormone was produced in TRANSPLASTOMIC tobacco leaves at levels exceeding 7% of the total soluble protein[14] and human serum albumin was produced at levels greater than 11% of the total soluble protein[15]. Higher expression levels have also been obtained in other plant species. For example, hirudin, which is expressed as a fusion to the oil-body protein oleosin, has been produced in transgenic canola at 0.3% of the total seed protein[16].

Recombinant antibodies. Antibodies are complex glycoproteins that recognize and bind to target antigens with great specificity. This individual and specific binding activity allows antibodies to be used for a range of applications, including the diagnosis, prevention and treatment of disease[17].

The production of antibodies in plants represents a special challenge because the molecules must fold and assemble correctly to recognize their cognate antigens. Typical serum antibodies are tetramers of two identical heavy chains and two identical light chains; however, there are more complex forms, such as secretory antibodies, which are dimers of the typical serum antibody and include two extra polypeptide chains. Two different cell types are required to assemble such antibodies in mammals, but plants that express four different transgenes can assemble these antibodies in a single cell[18].

As well as full size immunoglobulins, further antibody derivatives have been expressed successfully in plants, including Fab fragments, scFvs, bispecific Fvs, DIABODIES, MINIBODIES, single variable domains, antibody-fusion proteins, large single-chain antibodies and camelid heavy-chain antibodies (BOX 1). Unlike the proteins discussed above, antibodies have been expressed in many different plant systems and extensive comparisons have been carried out to determine optimal parameters for expression. Therefore, yields in excess of 1% of the total soluble protein are routine and antibodies are set to become the first generation of plant-derived therapeutic proteins to be produced on a commercial basis.

Although tobacco was used in the early studies, many antibodies are now produced in cereal seeds instead, as protein accumulation in dry seeds allows long-term storage at ambient temperatures without notable degradation or loss of activity. Antibodies have also been produced in less established plant-based expression platforms such as AGROINFILTRATED LEAVES, cell-suspension cultures and virus-infected plants[4]. A particular advantage of these systems is the short development time and rapid onset of protein production, which can be counted in weeks rather than the several months that are required to establish a production line of transgenic or transplastomic plants.

At least six types of plant-derived recombinant antibody have progressed to the preclinical testing stage, with the most advanced product now undergoing phase II clinical trials. These front-runners are discussed briefly in BOX 2.

Recombinant subunit vaccines. Since the HBV vaccine was produced and tested (see above), the concept of oral vaccination with raw fruit, vegetables, leaves and seeds has risen in popularity. Edible plants, rather than tobacco, are now the focus of research into HBV vaccine production in plants[19], and clinical trials have been carried out with the surface antigen that is expressed in potato and lettuce[20]. Two further vaccine candidates have reached the clinical trials stage, both of which are

ENTEROTOXIGENIC
Producing toxins in the gut that specifically affect the intestinal mucosa.

expressed in potato: the heat-labile toxin B subunit (LT-B) of ENTEROTOXIGENIC *Escherichia coli* (ETEC) and the capsid protein of Norwalk virus (NVCP)[21,22]. These antigens, from two important enteric pathogens, might be ideal oral-subunit vaccine candidates, as both are multimeric structures that survive in the extreme conditions of the human gut. Each protein accumulated to high levels in potato tubers and was correctly assembled into oligomers. Clinical trials with the LT-B vaccine showed that the consumption of raw potato tubers that contained 0.3–10 mg of LT-B produced high titres of mucosal and systemic antibodies[21].

Other proteins of medical relevance. Plants have been used to produce several other proteins with direct or indirect medical applications. These include the milk proteins β-casein and lysozyme, which could be used to improve child health[12,23], and protein polymers that

Table 2 | **Important pharmaceutical proteins that have been produced in plants**

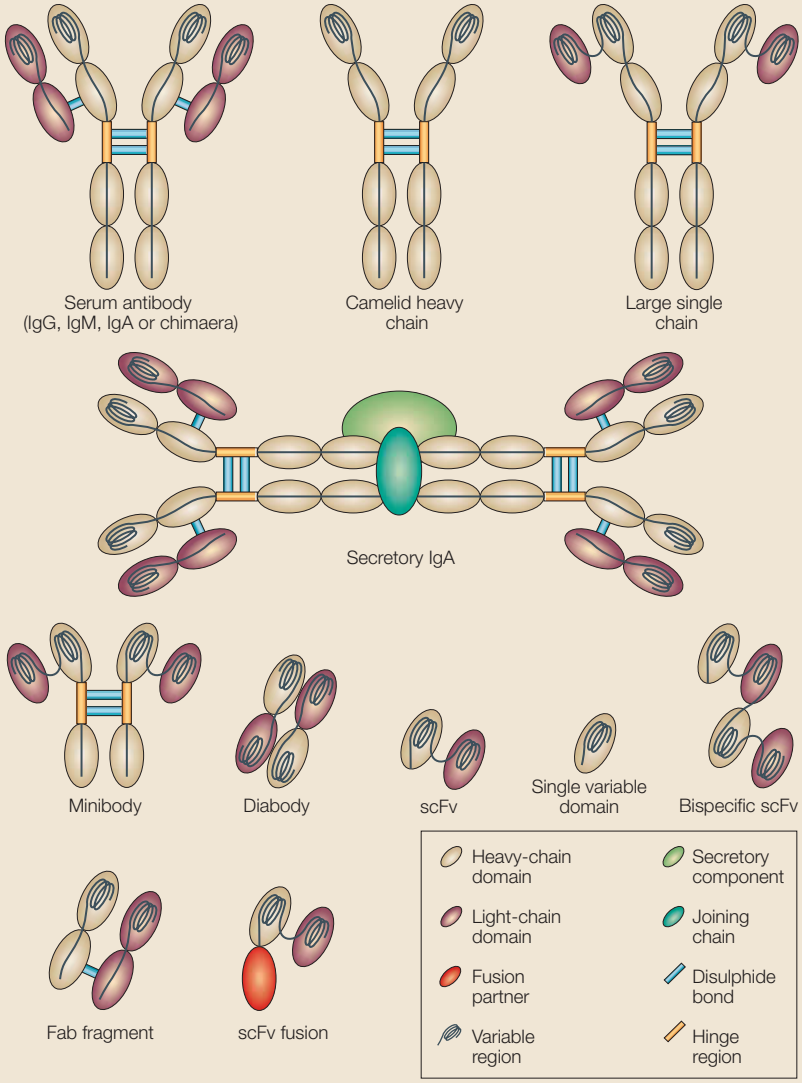
Protein	Host plant system	Comments	References
Human biopharmaceuticals			
Growth hormone	Tobacco, sunflower	First human protein expressed in plants; initially expressed as fusion protein with *nos* gene in transgenic tobacco; later the first human protein expressed in chloroplasts, with expression levels ~7% of total leaf protein	6,14
Human serum albumin	Tobacco, potato	First full size native human protein expressed in plants; low expression levels in transgenics (0.1% of total soluble protein) but high levels (11% of total leaf protein) in transformed chloroplasts	15,98
α-interferon	Rice, turnip	First human pharmaceutical protein produced in rice	99
Erythropoietin	Tobacco	First human protein produced in tobacco suspension cells	100
Human-secreted alkaline phosphatase	Tobacco	Produced by secretion from roots and leaves	59,60
Aprotinin	Maize	Production of a human pharmaceutical protein in maize	101
Collagen	Tobacco	First production of human structural-protein polymer; correct modification achieved by co-transformation with modification enzyme	13,26
α1-antitrypsin	Rice	First use of rice suspension cells for molecular farming (see REF. 102 for discussion of antibody production in rice cell culture)	103
Recombinant antibodies			
IgG1 (phosphonate ester)	Tobacco	First antibody expressed in plants; full length serum IgG produced by crossing plants that expressed heavy and light chains	7
IgM (neuropeptide hapten)	Tobacco	First IgM expressed in plants and protein targeted to chloroplast for accumulation	104
SIgA/G (*Streptococcus mutans* adhesin)	Tobacco	First secretory antibody expressed in plants; achieved by sequential crossing of four lines carrying individual components; at present the most advanced plant-derived pharmaceutical protein	89,90, 105
scFv-bryodin 1 immunotoxin (CD 40)	Tobacco	First pharmaceutical scFv produced in plants; first antibody produced in cell-suspension culture	106
IgG (HSV)	Soybean	First pharmaceutical protein produced in soybean	72
LSC (HSV)	*Chlamydomonas reinhardtii*	First example of molecular farming in algae	107
Recombinant subunit vaccines			
Hepatitis B virus envelope protein	Tobacco	First vaccine candidate expressed in plants; third plant-derived vaccine to reach clinical trials stage	8,19,20
Rabies virus glycoprotein	Tomato	First example of an 'edible vaccine' expressed in edible plant tissue	77
Escherichia coli heat-labile enterotoxin	Tobacco, potato	First plant vaccine to reach clinical trials stage	21,108
Norwalk virus capsid protein	Potato	Second plant vaccine to reach clinical trials stage	22
Diabetes autoantigen	Tobacco, potato	First plant-derived vaccine for an autoimmune disease	109
Cholera toxin B subunit	Tobacco, potato	First vaccine candidate expressed in chloroplasts	65
Cholera toxin B and A2 subunits, rotavirus enterotoxin and enterotoxigenic *E. coli* fimbrial antigen fusions	Potato	First plant-derived multivalent recombinant antigen designed for protection against several enteric diseases	110
Porcine transmissible gastroenteritis virus glycoprotein S	Tobacco, maize	First example of oral feeding inducing protection in an animal	111

HSV, herpes simplex virus; IgG, immunoglobulin G; IgM, immunoglobulin M; LSC, long single chain; *nos*, nopaline synthase; scFv, single-chain FV fragment; SIgA, secretory immunoglobulin A.

Box 1 | **Recombinant antibodies that are expressed in plants**

Plants have been used to express many different antibody forms, including full-length immunoglobulins and artificially constructed derivatives[86,87]. Several different types of immunoglobulin (Ig) have been produced successfully in plants, including IgG (various subclasses), IgA and a chimeric IgA/G. These have ranged in sequence and structure from completely murine to fully humanized. All these molecules comprise two identical heavy chains and two identical light chains, with each chain encoded by a separate transgene. Typically, the genes are introduced into separate plant lines and then stacked by creating hybrids of the transgenic parents. However, multiple gene transfer can achieve the same aim in one transformation procedure, and the correct assembly of immunoglobulins has also been shown after co-infection of tobacco plants with two viruses that each express a separate chain[88]. A variation on this theme is the camelid heavy-chain antibody, which lacks a light-chain component and can therefore be expressed as a single transgene. Secretory IgAs are dimers of the typical serum-type immunoglobulins and include two extra components: the secretory component and the joining chain. Four separate transgenes are required to produce such molecules[89].

Smaller engineered antibody derivatives, which for convenience have both the heavy- and light-chain variable regions on the same polypeptide chain, have also been expressed in plants[86,87]. These include derivatives that form spontaneous dimers (minibodies, diabodies and large single chains) and those that bind univalently to the antigen (single-chain Fv fragments (scFvs) and single variable regions). More specialized derivatives include bispecific scFvs, which contain the variable regions from two parent immunoglobulins and recognize two unrelated antigens, and scFv fusion proteins in which the scFv is genetically fused to a toxin, cytokine or enzyme.



could be used in surgery and tissue replacement. Early experiments with artificial polymers that were based on bovine elastin provided disappointing yields, even though mRNA levels were high, which indicated inefficient protein synthesis[24,25]. More recently, it has been shown that human collagen can be produced in transgenic tobacco plants and that the protein is spontaneously processed and assembled into its typical triple-helical conformation[13]. The original plant-derived collagen had a low thermal stability owing to the lack of hydroxyproline residues, but this was remedied by co-expressing the enzyme proline-4-hydroxylase[26]. A synthetic spider silk has also been expressed in transgenic plants[27]. Genes that were modelled on the endogenous silk protein genes of the spider *Nephila clavipes* were synthesized in the laboratory and introduced into tobacco and potato. Proteins up to 100 kDa in size and with 90% identity to the genuine silk protein were produced in tobacco leaves, potato leaves and potato tubers, at up to 2% of the total soluble protein.

DICOLYLEDONOUS PLANTS (Dicots). Broad-leaf flowering plants the seeds of which contain two cotyledons (embryonic seed leaves that either remain in the seed when the plant germinates or emerge and become green). Examples include potato, tomato, tobacco and all peas and beans.

MONOCOTYLEDONOUS PLANTS Narrow-leaf plants the seeds of which contain one cotyledon. Examples include cereals, grasses, orchids and lilies.

Genetic aspects of molecular farming in plants

Gene constructs. One of the main aims in molecular farming is the production of recombinant proteins at high yields. To achieve high yields, expression-construct design must optimize all stages of gene expression, from transcription to protein stability. Expression constructs are chimeric structures in which the transgene is bracketed by various regulatory elements that are known to be active in plants. For high-level transcription, the two most important elements are the promoter and the polyadenylation site, which are often derived from the 19S and 35S transcripts of the cauliflower mosaic virus (CaMV)[28,29]. The CaMV 35S promoter is now the most popular choice in DICOTYLEDONOUS PLANTS (dicots). It is a strong constitutive promoter that can be made even more active by duplicating the enhancer region[30]. However, this promoter has a lower activity in MONOCOTYLEDONOUS PLANTS (monocots), so alternatives such as the maize ubiquitin-1 promoter are preferred[31]. The presence of an intron in the 5′ untranslated region of the expression construct has also been

Box 2 | **Plant-derived antibodies: the front-runners**

Many antibodies and antibody derivatives have been produced in plant systems, but only seven have reached advanced stages of product development. Some of these will probably represent the first plant-derived recombinant pharmaceutical proteins to achieve commercial status.

Avicidin

This is a full size immunoglobulin G (IgG) that recognizes the anti-epithelial cellular-adhesion molecule (EpCAM) antigen, which is a marker of colorectal cancer. The antibody has been produced in both animal and plant systems, and the two variants behave in the same way, despite some biochemical differences. The plant-derived antibody, which is produced in corn, was the first to be administered to humans and showed some anti-cancer activity. However, the molecule was withdrawn completely by the developers (NeoRx and Monsanto) because of diarrhoea and other side effects in phase II trial patients, which were probably caused by cross-reaction with related EPITOPES on the intestinal epithelium. These effects were not specific to the plant-derived antibody.

CaroRx

CaroRx is a chimeric secretory IgA/G that is produced in transgenic tobacco plants through the expression of four separate transgenes. These were stacked by the sequential crossing of independent transgenic plants that each expressed a different component[89,90]. The antibody recognizes the main adhesion protein of *Streptococcus mutans*, which is the oral pathogen that is responsible for tooth decay in humans. Phase II clinical trials have shown that topical application, after bacteria have been removed from the mouth, helps to prevent recolonization by *S. mutans* for several months.

T84.66

T84.66 is a monoclonal antibody that recognizes carcinoembryonic antigen (CEA), which is a well-characterized marker of epithelial cancers (carcinomas). It has been widely tested for cancer imaging and therapy. T84.66 and recombinant derivatives such as single-chain Fv fragment (scFv) 84.66 have been produced in many different plant systems and used as objective comparators for expression efficiency and product quality[33,68,74,91]. A fusion protein that combines scFv84.66 and interleukin-2 has also been produced and could be used to stimulate lymphokine-activated killer cells and tumour-infiltrating lymphocytes near tumour masses.

Anti-HSV and anti-RSV

A full-length humanized IgG1 that recognizes herpes simplex virus (HSV)-2 glycoprotein B has been expressed in transgenic soybean and Chinese hamster ovary (CHO) cells[72]. Antibodies from both sources have been shown to prevent vaginal HSV-2 transmission in mice after topical application and would represent an inexpensive preventative for sexually transmitted diseases if they worked similarly in humans. This antibody, along with an IgG that recognizes the R9 protein of respiratory syncitial virus, is being developed by the United States biotechnology company Epicyte Pharmaceutical.

38C13

This scFv antibody is based on the idiotype of malignant B lymphocytes in the well-characterized mouse lymphoma cell line 38C13. Administration of the antibody to mice resulted in the production of anti-idiotype antibodies that are able to recognize 38C13 cells, which help to protect the mice against a lethal challenge with injected lymphoma cells[88]. This system could be adapted to produce antibodies that recognize unique markers on the surface of any malignant B cell and could therefore be an effective therapy for human diseases such as non-Hodgkin lymphoma. The antibodies were produced using virus-infected plants rather than transgenic plants, which is a strategy that is well suited to the rapid and small-scale production that is required to treat individual patients with unique antibodies. Large Scale Biology Corp. has completed phase I trials.

PIPP

PIPP is a monoclonal antibody that recognizes human chorionic gonadotropin (hCG). The full-length monoclonal antibody, as well as scFv and diabody derivatives, has been produced in transgenic plants and by agroinfiltration in transiently transformed tobacco[92]. Each of the antibodies was able to inhibit the hCG-stimulated production of testosterone in cultured LEYDIG CELLS and to delay uterine weight gain in mice, which are the standard tests for hCG activity. The antibodies could be used for the diagnosis and/or therapy of tumours that produce hCG, pregnancy detection and (emergency) contraception.

EPITOPE
A single antigenic determinant on a protein that is recognized by an antibody. A single protein can have many epitopes.

LEYDIG CELLS
Interstitial cells in the testis that are responsible for the production of male sex hormones, such as testosterone, and are important in male sexual differentiation.

shown to enhance transcription in monocots[32]. Widely used polyadenylation sites include those from the CaMV 35S transcript, the *Agrobacterium tumefaciens nos* gene and the pea *ssu* gene.

Promoters that allow the expression of a transgene in a particular environmental, developmental or tissue-specific manner might also be useful. For example, there are several advantages to the restriction of transgene expression to cereal seeds[33] and potato tubers[34] using tissue-specific promoters, such as those from the maize zein, rice glutelin, wheat glutenin and pea legumin genes. The advantages of such promoters include the increased stability of the protein and the avoidance of protein accumulation in vegetative organs, so preventing toxicity to the host plant and contact with non-target organisms. Inducible promoter systems, which respond to external chemical[35,36] and physical[37] stimuli, might also be used to restrict transgene expression on a temporal basis.

For example, the mechanical gene activation (MeGA) system that was developed by Cramer (CropTech Corp., Virginia, United States) uses a tomato hydroxy-3-methylglutaryl CoA reductase 2 (HMGR2) promoter, which is inducible by mechanical stress. Transgene expression is activated when harvested tobacco leaves are sheared during processing, which leads to the rapid induction of protein expression, usually within 24 hours. Many other inducible promoters have been developed — for example, those that use ethanol, dexamethasone and the insecticide methoxyfenozide — and have recently been reviewed[35].

Transgenes from heterologous species often have a different codon bias to the host plant, which might result in pausing at disfavoured codons and truncation, misincorporation or frameshifting. Such effects can be avoided by introducing silent mutations into the coding region of the transgene by SITE-DIRECTED MUTAGENESIS, which brings transgene codon usage in line with that of the host.

One of the most important factors governing the yield of recombinant proteins is subcellular targeting, which affects the interlinked processes of folding, assembly and post-translational modification. Comparative experiments with recombinant antibodies have shown that the secretory pathway is a more suitable environment for folding and assembly than the cytosol, leading to higher yields[38]. Proteins are targeted to the secretory pathway through the inclusion of an N-terminal SIGNAL PEPTIDE in the expression construct. Although most antibodies accumulate to higher levels in the secretory pathway compared with the cytosol, there are some notable exceptions[39,40], which indicate that intrinsic features of each antibody might also influence their overall stability. The oxidizing environment of the endoplasmic reticulum (ER), the lack of proteases and the abundance of MOLECULAR CHAPERONES are important factors for correct protein folding and assembly. Also, protein glycosylation occurs only in the endomembrane system and this modification is required for the correct function of many proteins of human origin.

In the absence of further targeting information, proteins in the endomembrane system are secreted to the APOPLAST, where they might be retained or secreted into the environment. However, antibody yields can be increased even further if the protein is retained in the ER lumen using an H/KDEL C-terminal tetrapeptide tag, as this compartment has a stabilizing influence[41]. Yields are generally twofold to tenfold greater for ER-retention compared with secretion. Proteins that are retained in this manner are not modified in the Golgi apparatus, which means that they have high-mannose glycans but no plant-associated xylose and fucose residues (see below). Although the measures discussed above will help to achieve high intrinsic yields, the actual amount of recombinant protein that is then obtained ultimately depends on the processing and purification methods (BOX 3).

Other factors that influence transgene expression. Transgene expression is influenced by several factors that cannot be controlled precisely through construct design, which lead to variable transgene expression and, in some cases, its complete inactivation[42]. Such factors include the position of transgene integration, the structure of the transgenic locus, gene-copy number and the presence of truncated or rearranged transgene copies. Several strategies have been adopted in an attempt to minimize variation in transgene expression, including, most recently, the use of viral genes that suppress gene silencing[43]. Preliminary studies have indicated that the co-transformation of plants with a primary transgene and a viral-silencing suppressor notably increases transgene expression level[44]. The ability to integrate single-copy transgenes into precise locations in the plant nucleus would eliminate POSITION EFFECTS and the problems that are associated with variable locus structure. Several laboratories are therefore investigating ways to improve the efficiency of gene targeting in plants[45]. In practice, however, commercially developed transgenic plants undergo an enormous amount of screening to identify phenotypic, yield and agronomic variation. The screening includes

SITE-DIRECTED MUTAGENESIS
An *in vitro* mutagenesis procedure that is often carried out using the polymerase chain reaction in which specific mutations are introduced into a DNA molecule.

SIGNAL PEPTIDE
A short sequence of mainly hydrophobic amino acids at the N-terminus of secreted proteins. This peptide is captured by a signal-recognition particle as it emerges from the ribosome, which allows the ribosome to be transported to the endoplasmic reticulum.

MOLECULAR CHAPERONES
Proteins the function of which is to ensure correct folding of other proteins during or after synthesis, or the refolding of denatured proteins.

APOPLAST
The extracellular space. In plants, this is a large and continuous network of cavities under the cell wall. Proteins that are secreted from the cell often remain trapped here.

POSITION EFFECTS
When transgenes integrate into genomic DNA, the expression level is often influenced by the surrounding chromatin. Local regulatory elements, such as enhancers, also influence transgene expression. Position effects lead to wide variations in transgene expression levels, even in plants that are transformed with identical constructs.

AFFINITY TAGS
Short peptide sequences added to recombinant proteins, which bind strongly to particular affinity matrices and can be used to purify recombinant proteins.

Box 3 | **Production issues**

Highly efficient purification schemes are a prerequisite for the conventional use of recombinant proteins as pharmaceuticals, and this requirement must be built into strategies for molecular farming in plants[93–95]. Although AFFINITY TAGS, such as His_6 or the FLAG epitope, can be used for purification, this type of modification alters the primary structure of the protein and might adversely affect its properties. So, it might be necessary to avoid affinity tags and devise specific purification schemes for individual proteins on the basis of their native structures. Protocols are available for the purification of antibodies from whole plants, plant cell-suspension cultures, leaves and seeds, using protein-A and protein-G-based affinity matrices[96], although this approach is not suitable for some sub-classes of antibodies and most small recombinant antibody derivatives (including scFvs). In general, the costs of processing are reduced when the product is more concentrated in the starting material and this is one of the advantages of expressing recombinant proteins in the seeds of transgenic plants, in which high levels of the product can accumulate in a small volume[33]. If conventional extraction from seeds is too expensive, further strategies to assist purification might be used. One example is the oleosin-fusion platform that was developed by SemBioSys Genetics Inc., in which the target recombinant protein is expressed in oilseed crops as a fusion with oleosin. The fusion protein can be recovered from oil bodies using a simple extraction procedure, and the recombinant protein separated from its fusion partner by endoprotease digestion[16]. Similarly, Schillberg and colleagues have devised a strategy in which recombinant proteins are expressed as fusion constructs that contain an integral membrane-spanning domain that is derived from the human T-cell receptor[97]. The recombinant protein accumulates at the plasma membrane and can be extracted in a small volume using appropriate buffers and detergents.

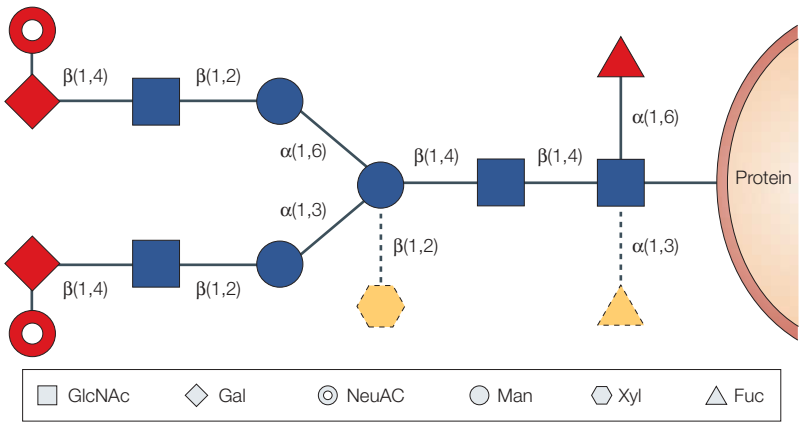


Figure 1 | **Complex long-chain glycan structure in plants and humans.** To 'humanize' the recombinant proteins that are made in plants, α(1,3) fucose and β(1,2) xylose residues must be removed (dotted lines), whereas galactose and sialic-acid residues must be added. Blue residues are common to plants and humans. Red residues are found in humans and not plants, so they need to be added. Yellow residues are found in plants but not humans and need to be removed. Fuc, fucose; Gal, galactose; GlcNAc, N-acetylglucosamine; Man, mannose; NeuAC, acetylneuraminic acid (sialic acid); Xyl, xylose.

identification of the site of transgene insertion, which allows a rational risk assessment to be made of the likelihood of adverse unintentional effects that result from the transformation process.

Transformation methods. Two general methods are used to generate transgenic plant lines for molecular farming: AGROBACTERIUM-MEDIATED TRANSFORMATION and particle bombardment, in which DNA-coated microprojectiles are accelerated into plant tissue. Each method has advantages and disadvantages, and the choice depends on a combination of factors, including the selected host species, local expertise and intellectual property issues. Other methods, such as WHISKER TRANSFORMATION, ELECTROPORATION and PROTOPLAST TRANSFORMATION, have not so far been used for molecular-farming applications.

The soil pathogen *A. tumefaciens* provides a simple method for the transformation of most dicot species and is commonly used for molecular farming in tobacco, alfalfa, pea, tomato and potato[44,46]. Monocots can also be transformed by *Agrobacterium*, but in most cases the technology has been optimized for selected model varieties. Particle bombardment shows less genotype dependence and might be the preferred transformation method for cereals, such as rice, wheat and maize, as well as soybean and other legumes[47]. Particle bombardment is also necessary for plastid transformation, as the *Agrobacterium* T-DNA complex is targeted to the nucleus and is therefore unsuitable for gene transfer to chloroplasts[44]. Transformation can also be achieved using *Agrobacterium rhizogenes*, but this organism is preferred for the production of transgenic root cultures.

These transformation methods generally lead to the introduction of superfluous DNA sequences into the nuclear genome[48]. In the case of *Agrobacterium*-mediated transformation, this is because inefficient processing of the T-DNA BORDER SEQUENCES often results in the co-transfer of flanking vector sequences that might sometimes correspond to the entire plasmid[49]. In the case of particle bombardment, superfluous DNA transfer occurs because whole plasmids are generally used to coat the microprojectiles[47]. Superfluous DNA transfer is a regulatory problem under the strict new guidelines for the release of genetically modified organisms into the environment. Therefore, several strategies have been developed to avoid the transfer of vector sequences during transformation (see also the article by Stewart *et al.* in this issue). Incorporation of the *barnase* gene outside the T-DNA border sequences is one approach that works during *Agrobacterium*-mediated transformation. This ensures that all plant cells that contain vector sequences linked to the T-DNA are killed, as *barnase* expression is lethal[50]. Clean-DNA techniques, in which only the necessary transgenes but no vector backbone or superfluous marker genes are introduced into the plant, were also developed for particle bombardment, after it was shown that the microprojectiles can be coated with minimal cassettes (essentially the promoter, transgene and polyadenylation site), without compromising transformation efficiency. Furthermore, transgenic loci in clean-DNA plants are considerably simpler than those of whole-plasmid transformants, and the plants show a notable reduction in the frequency of transgene silencing[51].

Post-translational modification. The protein-synthesis pathway is highly conserved between plants and animals, so human transgenes that are expressed in plants yield proteins with identical amino-acid sequences to their native counterparts. However, there are some important differences in post-translational modification. One example, as discussed above, concerns the inability of transgenic plants to correctly modify human collagen unless a gene that encodes proline-4-hydroxylase is also expressed[26]. The main difference between proteins that are produced in animals and plants, however, concerns the synthesis of glycan side chains. All eukaryotes add glycan chains to proteins as they pass through the secretory pathway, but owing to differences in the levels of different modification enzymes, the glycan-chain structures vary widely across different taxa. Plant-derived recombinant proteins tend to lack the terminal galactose and sialic-acid residues that are normally found in mammals, but have the carbohydrate group α(1,3)fucose, which has a (1,6) linkage in animal cells, and β(1,2)xylose, which is absent in mammals although present in invertebrates (FIG. 1).

These minor differences in glycan structure could potentially change the activity, biodistribution and longevity of recombinant proteins compared with the native forms. The possibility of plant-specific glycans inducing allergic responses in humans has been considered[52] and the finding that human serum contains antibodies that are reactive against these residues has been interpreted as evidence that the α(1,3)fucose and β(1,2)xylose residues might lead to adverse reactions[53]. However, carbohydrate epitopes are rarely allergenic. Moreover, the presence of antibodies in serum is not indicative of an adverse reaction. Finally, these glycan residues are also associated with every normal plant

AGROBACTERIUM-MEDIATED TRANSFORMATION
Transformation that is achieved using the natural gene-transfer mechanism of *Agrobacterium tumefaciens*.

WHISKER TRANSFORMATION
Transformation that is achieved by mixing walled plant cells with silicon carbide fibres that penetrate the cell wall and membrane, which generate pores through which DNA can be taken up into the cell.

ELECTROPORATION
Transformation that is achieved by exposing cells or protoplasts to a brief pulse of electricity, which results in the formation of transient membrane pores through which DNA can be taken up into the cell.

PROTOPLAST TRANSFORMATION
Any technique for introducing DNA into unwalled plant cells (protoplasts), such as calcium phosphate transfection, PEG transfection or electroporation.

T-DNA BORDER SEQUENCES
Imperfect 25 bp direct repeat sequences that flank the piece of DNA that is transferred to the plant genome by *Agrobacterium tumefaciens*. These sequences are recognized by the bacterial VIRD1 and VIRD2 proteins, which form an endonuclease complex. Cleavage of the border sequences initiates T-DNA transfer.

glycoprotein that is found in our diet. So, it is highly unlikely that they will be associated with adverse reactions. Indeed, studies in which mice were a dministered a recombinant antibody that contained plant-specific glycans showed no evidence of an anti-glycan immune reaction[54]. Nevertheless, the perceived negative effect of 'foreign' glycan structures is one of the most important issues that affect the use and acceptance of plant-derived recombinant proteins. Therefore, recent attention has focused on the development of strategies to 'humanize' the glycosylation patterns of recombinant proteins. Warner[55] provides an overview of the biochemistry of glycan-chain synthesis in different expression hosts, and lists the changes that are required to produce proteins with typical human glycan structures in plants. In the moss *Physcomitrella patens*, gene targeting has been used to disable the plant-specific fucosyltransferase and xylosyltransferase enzymes. Strategies that have been attempted in transgenic plants include the use of purified human β(1,4)-galactosyltransferase and sialyltransferase enzymes to modify plant-derived recombinant proteins *in vitro*[56], and the expression of human β(1,4)-galactosyltransferase in transgenic tobacco plants to produce recombinant antibodies with galactose-extended glycans[57]. In the latter case, ~30% of the recovered antibody was galactosylated. This is similar to the proportion of galactosylated antibodies that are produced by HYBRIDOMA CELLS. *In vivo* sialylation is unlikely to be achieved in the near future because plants seem to lack the metabolic pathway for the precursors of sialic acid, so several new enzymes would need to be introduced and coordinately expressed.

To place the issue in perspective, it should be remembered that there is natural variation in glycan structures, with many proteins having several glycosylation sites and, even in mammalian cells, a range of glycoforms. There are recognizable differences in glycan structure even when comparing native human proteins to those produced in rodent cell lines. For example, human antibodies contain only the sialic-acid residue N-acetylneuraminic acid (NANA), whereas rodents produce a mixture of NANA and N-glycosylneuraminic acid (NGNA)[58]. At this stage, it is hard to generalize about how crucial the humanized glycosylation of plant-derived pharmaceuticals is, and whether it might be more important for some classes of proteins than for others. Similarly, there are still not enough data to address whether the method of administration of recombinant proteins (oral versus injection) could make a difference in terms of the immune response that might occur.

HYBRIDOMA CELLS
A hybrid cell line that is created by fusing a mortal antibody-producing B-lymphocyte with an immortalized myeloma line. The hybridoma line is immortal and produces a continuous supply of a particular monoclonal antibody.

SOLANACEAE
A family of flowering plants (order Solanales) that comprise ~100 genera and ~2,500 species, many of which are economically important as food or medicinal crops. Examples include tobacco, potato and tomato.

Plant-expression hosts

Tobacco production systems. Tobacco has an established history as a model system for molecular farming and is the most widely used species for the production of recombinant pharmaceutical proteins at the research-laboratory level[4,37]. The main advantages of tobacco include the mature technology for gene transfer and expression, the high biomass yield (more than 100,000 kg per hectare for close-cropped tobacco), the potential for rapid scale-up owing to prolific seed production, and the availability of large-scale infrastructure for processing. Although many tobacco cultivars produce high levels of toxic alkaloids, there are low-alkaloid varieties that can be used for the production of pharmaceutical proteins[4].

In most cases, nuclear transgenic plants have been used for production and the proteins have been extracted from leaf tissue. Targeting proteins to the secretory pathway in tobacco can result in them being exuded from the roots or leaves (rhizosecretion and phyllosecretion, respectively)[59,60]. Although not widely adopted so far, this strategy is potentially useful because no cropping or harvesting is necessary. The technology is under commercial development for the production of human secreted alkaline phosphatase (see Phytomedics Inc. in online links box). Perhaps surprisingly, even large molecules can be rhizosecreted from transgenic plants. For example, a monoclonal antibody was secreted into hydroponic culture medium resulting in a yield of 11.7 μg antibody per gram of dry root mass per day[61]. Plants can reasonably be expected to survive in a hydroponic system for many months, and proteins are relatively easily purified from culture medium compared with extraction from leaves, so rhizosecretion represents an attractive option for antibody production.

As an alternative to nuclear transgenics, transplastomic plants are produced by introducing DNA into the chloroplast genome rather than the nuclear genome, a process that is generally achieved by particle bombardment[62,63]. The advantages of chloroplast transformation include the high transgene-copy number (there can be several thousand chloroplasts in a photosynthetic cell) and the absence of position effects and transgene silencing. In combination, these properties can lead to astonishing levels of expression, in the best cases exceeding 25% of the total soluble protein[64]. Further advantages of chloroplast engineering include the ability to express several genes as operons and the accumulation of recombinant proteins in the chloroplast (which reduces toxicity to the host plant). As discussed above, both human growth hormone and serum albumin have been produced at high levels in tobacco chloroplasts, and each protein was found to be structurally authentic and biologically active[14,15]. More recently, a tetanus toxin fragment has been expressed in tobacco chloroplasts and was shown to induce protective levels of anti-tetanus antibodies[64]. The cholera toxin B subunit has also been expressed in chloroplasts, which shows that plastids can fold and assemble oligomeric proteins correctly[65]. One disadvantage of the chloroplast transgenic system is that plastids do not carry out glycosylation. It is therefore unlikely that chloroplasts could be used to synthesize human glycoproteins in cases in which the glycan-chain structure is crucial for protein activity. Another limitation is that chloroplast transformation outside the SOLANACEAE still presents a formidable technical challenge[66].

Recombinant proteins can also be produced in plant cell cultures. Tobacco suspension cells are generated by

the continuous agitation of FRIABLE CALLUS TISSUE, which results in a homogeneous suspension of single cells and small clumps. The cultures can be maintained in conventional microbial fermenters with only minor technical modifications, and various different culture modes can be used, including BATCH, FED-BATCH, PERFUSION AND CONTINUOUS FERMENTATION. Fischer and colleagues have reported the expression of several recombinant proteins, including several antibody derivatives, in a suspension cell line that was derived from the tobacco strain BY-2 (REF. 67). This approach is particularly advantageous when defined and sterile production conditions are required together with straightforward purification protocols. Recombinant proteins that are secreted into the culture medium are more easily purified, although proteins that are larger than 20–30 kDa tend to be retained in the apoplast and must be released by mechanical or enzymatic disruption.

Cereals and legumes. One of the disadvantages of recombinant-protein production in tobacco is the instability of the product, which means that the leaf tissue must be frozen or dried for transport, or processed at the farm. By contrast, the accumulation of recombinant antibodies in seeds allows long-term storage at ambient temperatures because the proteins amass in a stable form. Seeds have the appropriate biochemical environment for protein accumulation, and achieve this through the creation of specialized storage compartments, such as protein bodies and storage vacuoles, which are derived from the secretory pathway. Seeds are also desiccated, which reduces the exposure of stored proteins to non-enzymatic hydrolysis and protease degradation. It has been shown that antibodies that are expressed in seeds remain stable for at least three years at room temperature with no detectable loss of activity[33]. Cereal seeds also lack the phenolic substances that are present in tobacco leaves, so increasing the efficiency of downstream processing. The main concern about the use of cereals and other established food crops for the production of pharmaceuticals relates to the potential for genes to spread into crops that are grown for food purposes, and the possibility of inadvertent contamination during seed collection and storage (see the article by Stewart *et al.* in this issue).

Important variables that must be considered when choosing a cereal production crop include the grain yield per hectare, the yield of recombinant protein per unit biomass, the ease of transformation and the speed of scale-up. The same single-chain variable-fragment antibody has been expressed in rice, wheat and tobacco to compare production levels in leaves and seeds[68]. With the optimal promoter system (the enhanced CaMV35S promoter for tobacco and the ubiquitin-1 promoter for rice) it was found that rice plants showed the highest yields per unit biomass, and levels were lowest in wheat. However, the wheat system is still under development and improved construct design will probably give rise to higher yields.

Maize is now the main commercial production crop for recombinant proteins, which reflects advantages such as high biomass yield, ease of transformation and *in vitro* manipulation, and ease of scale-up. These factors, as well as intellectual property issues, prompted Prodigene to choose maize as the crop for the first commercial molecular-farming venture, which involved the production of the technical proteins avidin and β-glucuronidase[11,69]. Maize is also being used for the production of recombinant antibodies[70] and further technical/pharmaceutical enzymes, such as laccase, trypsin and aprotinin[71].

Alfalfa and soybean are legumes. They each produce lower amounts of leaf biomass than tobacco, but have the advantage of using atmospheric nitrogen through nitrogen fixation, thereby reducing the need for chemical inputs. Both species have been used to produce recombinant antibodies[72,73]. One of the potential advantages of alfalfa is the recent finding that recombinant antibodies are produced as a single glycoform rather than the heterogeneous collection of different glycoforms that is found in other plant systems (see Medicago in online links box). Grain legumes are also useful production crops because of the high protein content in the seeds. For example, pea is being developed as a production system, although at present the yields that are possible with this species are low[74].

Fruit and vegetables. The main benefit of fruit, vegetable and leafy salad crops is that they can be consumed raw or partially processed, which makes them particularly suitable for the production of recombinant subunit vaccines, food additives and antibodies for topical passive immunotherapy. As discussed above, potatoes have been widely used for the production of plant-derived vaccines and have been administered to humans in most of the clinical trials carried out so far. The potential of potato tubers for antibody production was first shown by Artsaenko and colleagues[34], and recently this crop has been investigated as a possible bulk-production system for antibodies[75]. Potatoes have also been used for the production of diagnostic antibody-fusion proteins[76] and human milk proteins[12,23]. Tomatoes, which were used to produce the first plant-derived rabies vaccine[77], are more palatable than potatoes and offer other advantages including high biomass yields (~68,000 kg per hectare) and the increased containment that is offered by growth in greenhouses. Lettuce is also being investigated as a production host for edible recombinant vaccines, and has been used in one series of clinical trials for a vaccine against HBV[20]. Bananas have been considered as hosts for the production of recombinant vaccines, as they are widely grown in the countries in which vaccines are most needed and can be consumed raw or as a puree by both adults and children[78].

Acceptability

Molecular farming involves the use of genetically enhanced plants to produce pharmaceuticals, and is therefore covered by a myriad of established and emerging regulations. Environmental biosafety issues, such as the potential for transgene spread and the possible

FRIABLE CALLUS TISSUE
Callus tissue is undifferentiated plant tissue, which grows when seeds or explants are cultured on media that contains an appropriate balance of plant hormones. Friable callus tissue is easily broken into fragments.

BATCH, FED-BATCH, PERFUSION AND CONTINUOUS FERMENTATION
Batch fermentation is a closed system in which all of the substrate is added at the beginning, whereas in the fed-batch process the substrate is added in increments as fermentation proceeds. Continuous fermentation is an open system in which substrate is added continuously at a steady rate. Perfusion fermentation is a continuous process that allows cells to be grown at high density, and so results in increased biomass and product yields.

toxicity of the recombinant proteins to non-target organisms, need to be addressed on a case-by-case basis depending on the location, production host and product, taking into account measures that have been used to reduce biosafety risks[79] (see the article by Stewart *et al.* in this issue). The rigorous scientific evaluation of each application for release will help in the development of effective risk-management strategies that will facilitate future decision-making processes. The main perceived risks are transgene spread by pollen dispersal, seed dispersal and horizontal gene transfer, and the effects of potentially toxic recombinant proteins on herbivores, pollinating insects and microorganisms in the RHIZOSPHERE. There is also concern that plant material that contains recombinant proteins could inadvertently enter the food chain.

Although it is apparent that pharmaceutical crops do not suffer the same acceptability problems as genetically modified food crops, risk assessment and environmental-impact studies must be carried out to the same level, to ensure the highest standards of responsibility and regulatory compliance. The removal of selectable marker genes[80] or the use of innocuous plant-derived markers for metabolic selection[81] is required to limit the incorporation of superfluous DNA sequences. The risk of transgene spread by pollen dispersal can be addressed by several physical and genetic barrier techniques, as well as by the choice of a suitable production crop that does not outcross with wild plants near the production site[82].

The risk of horizontal gene transfer is thought to be extremely low. The potential for the horizontal spread of DNA from transgenic plants to bacterial populations has been considered. No such gene flow has been shown with nuclear transgenic plants, despite many attempts to create conditions in the laboratory that would simulate such an occurrence. Recently, Kay *et al.*[83] showed the horizontal transfer of an antibiotic-resistance marker from the chloroplasts of transplastomic tobacco plants to opportunistic strains of *Acinetobacter* spp. However, transfer was achieved only under idealized conditions in which the bacteria were modified to contain a sequence that was homologous to the transgene. No transfer was observed to wild-type bacteria. The potential for transfer is not surprising in light of the prokaryotic origin of plastids, but the transferred gene would need to offer a selective advantage to the recipient bacteria to persist in natural populations. All plants, both natural and transgenic, have been shown to be covered with antibiotic-resistant bacteria, and these are a much more likely source of resistance genes that could jump to human pathogens[84].

The possible negative effect of recombinant proteins on non-target organisms can be addressed by the use of regulated promoters to restrict transgene expression to particular organs (for example, seeds) or to induce protein expression at particular times. The retrieval of proteins to the ER lumen, or the direction of proteins to other compartments such as the vacuole or chloroplast, can also improve the containment of proteins by preventing secretion to the apoplast and possible leaching into the LEAF GUTTATION FLUID or the root exudates. Another strategy is to express the recombinant protein as an inactive precursor that must be proteolytically cleaved before it shows biological activity.

Recently publicized incidents in which genetically enhanced crops have been inadvertently mixed with those destined for human consumption have highlighted the need for mechanisms to ensure the segregation of plants that express pharmaceuticals (for example, see Pew Initiative on Food and Biotechnology in online links box). Although greenhouse containment would address many of these biosafety issues, this negates the advantages of the technology in terms of low cost and large-scale production. If the main goal is to provide affordable medicines to developing countries, then these are important considerations for regulatory authorities. In any case, if food and feed crops are chosen for molecular farming, many other levels of safety can be built into the system. These include geographical isolation, differential planting seasons, the use of male-sterile plants and transplastomic plants, and the use of inducible promoters, as discussed above. It would also be helpful to use identity-preserved varieties, such as white tomatoes or maize, which are easily identified by their pigmentation. Identity preservation and tracking are important parts of the regulatory procedure for the production of pharmaceuticals in transgenic plants[85].

Conclusions

Plants have many advantages compared with traditional systems for the molecular farming of pharmaceutical proteins. These include the low cost of production, rapid scalability, the absence of human pathogens and the ability to fold and assemble complex proteins accurately. The commercial success of plant-derived avidin and β-glucuronidase, which are both available from Sigma Inc. at a lower cost than the native proteins, shows the potential of molecular farming in plants to compete in an already established market. Other important examples that show the potential of the technology include the synthesis, folding and assembly of multimeric proteins, such as serum and secretory antibodies, and subunit vaccines, which need to assemble into multimeric complexes to raise an immune response. Plants might one day surpass other production systems, particularly cultured animal cells, for the production of pharmaceutical proteins because of these economic and safety benefits. Several challenges remain to be met in terms of increasing yields, improving glycoprotein authenticity, removing processing bottlenecks and addressing biosafety and acceptability issues, as well as industry inertia. The difficulties that must be overcome as part of the existing drug regulatory process are not the least of these challenges; however, this complex issue is beyond the scope of this review. If these hurdles can be overcome, we might soon be in the position to produce any pharmaceutical protein on a scale that meets market demand. Ultimately, it might be possible to make protein-based pharmaceuticals available to everyone who needs them, at a cost that everyone can afford.

RHIZOSPHERE
The soil zone that surrounds plant roots, which is rich in microorganisms and in which interactions occur between plants and microbes.

LEAF GUTTATION FLUID
Fluid that seeps from the apoplast onto the leaf surface. In plants with large leaves, such as tobacco, large amounts of guttation fluid can be produced each day.

1. Schwartz, J. R. Advances in *Escherichia coli* production of therapeutic proteins. *Curr. Opin. Biotechnol.* **12**, 195–201 (2001).
2. Chu, L. & Robinson, D. K. Industrial choices for protein production by large-scale cell culture. *Curr. Opin. Biotechnol*. **12**, 180–187 (2001).
3. Houdebaine, L. M. Transgenic animal bioreactors. *Transgenic Res*. **9**, 305–320 (2000).
4. Fischer, R. & Emans, N. Molecular farming of pharmaceutical proteins. *Transgenic Res.* **9**, 279–299 (2000).
5. Giddings, G. Transgenic plants as protein factories. *Curr. Opin. Biotechnol*. **12**, 450–454 (2001).
6. Barta, A. *et al.* The expression of a nopaline synthase human growth hormone chimaeric gene in transformed tobacco and sunflower callus tissue. *Plant Mol. Biol*. **6**, 347–357 (1986).
7. Hiatt, A., Cafferkey, R. & Bowdish, K. Production of antibodies in transgenic plants. *Nature* **342**, 76–78 (1989).
8. Mason, H. S., Lam, D. M. K. & Arntzen, C. J. Expression of hepatitis B surface antigen in transgenic plants. *Proc. Natl Acad. Sci. USA* **89**, 11745–11749 (1992).
9. Thanavala, Y., Yang, Y.-F., Lyons, P., Mason, H. S. & Arntzen, C. J. Immunogenicity of transgenic plant-derived hepatitis B surface antigen. *Proc. Natl Acad. Sci. USA* **92**, 3358–3361 (1995).
10. Hood, E. E. *et al*. Criteria for high-level expression of a fungal laccase gene in transgenic maize. *Plant Biotechnol. J*. **1**, 129–140 (2003).
11. Hood, E. E. *et al*. Commercial production of avidin from transgenic maize: characterization of transformant, production, processing, extraction and purification. *Mol. Breeding* **3**, 291–306 (1997).
12. Chong, D. K. X. *et al*. Expression of the human milk protein β-casein in transgenic potato plants. *Transgenic Res*. **6**, 289–296 (1997).
13. Ruggiero, F. *et al*. Triple helix assembly and processing of human collagen produced in transgenic tobacco plants. *FEBS Lett*. **469**, 132–136 (2000).
14. Staub, J. M. *et al.* High-yield production of a human therapeutic protein in tobacco chloroplasts. *Nature Biotechnol*. **18**, 333–338 (2000).
This report shows that the secreted human protein somatotropin is soluble and biologically active when expressed in tobacco chloroplasts and has correctly formed disulphide bonds.
15. Fernandez-San Millan, A., Mingo-Castel, A., Miller, M. & Daniell, H. A chloroplast transgenic approach to hyper-express and purify human serum albumin, a protein highly susceptible to proteolytic degradation. *Plant Biotechnol*. **1**, 77–79 (2003).
16. Moloney, M., Boothe, J. & Van Rooijen, G. Oil bodies and associated proteins as affinity matrices. US Patent 6,509,453 (2003).
17. Chadd, H. E. & Chamow, S. M. Therapeutic antibody expression technology. *Curr. Opin. Biotechnol*. **12**, 188–194 (2001).
18. Ma, J. K. *et al*. Generation and assembly of secretory antibodies in plants. *Science* **268**, 716–719 (1995).
This study shows for the first time that secretory antibodies, with 10 polypeptide chains that represent the products of four genes, can be assembled correctly in transgenic plants. Two rounds of crossing, which involved four singly-transgenic lines, were required to stack all four transgenes in the same plant.
19. Richter, L. J., Thanavala, Y., Arntzen, C. J. & Mason, H. S. Production of hepatitis B surface antigen in transgenic plants for oral immunization. *Nature Biotechnol*. **18**, 1167–1171 (2000).
20. Kapusta, J. *et al*. A plant-derived edible vaccine against hepatitis B virus. *FASEB J.* **13**, 1796–1799 (1999).
21. Tacket, C. O. *et al.* Immunogenicity in humans of a recombinant bacterial-antigen delivered in a transgenic potato. *Nature Med*. **4**, 607–609 (1998).
22. Tacket, C. O. *et al.* Human immune responses to a novel Norwalk virus vaccine delivered in transgenic potatoes. *J. Infect. Dis*. **182**, 302–305 (2000).
23. Chong, D. K. X. & Langridge, W. H. R. Expression of full-length bioactive antimicrobial human lactoferrin in potato plants. *Transgenic Res*. **9**, 71–78 (2000).
24. Zhang, X., Urry, D. W. & Daniell, H. Expression of an environmentally friendly synthetic protein-based polymer in transgenic tobacco plants. *Plant Cell Reps*. **16**, 174–179 (1996).
25. Guda, C., Lee, S. B. & Daniell, H. Stable expression of a biodegradable protein-based polymer in stable tobacco chloroplasts. *Plant Cell Reps*. **19**, 257–262 (2000).
26. Merle, C. *et al*. Hydroxylated human homotrimeric collagen I in *Agrobacterium tumefaciens*-mediated transient expression and in transgenic tobacco plant. *FEBS Lett*. **515**, 114–118 (2002).
27. Scheller, J., Guhrs, K. H., Grosse, F. & Conrad, U. Production of spider silk proteins in tobacco and potato. *Nature Biotechnol*. **19**, 573–577 (2001).
28. O'Dell, J. T., Nagy, F. & Chua, N. H. Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. *Nature* **313**, 810–812 (1985).
29. Lawton, M. A. *et al*. Expression of a soybean β-conclycinin gene under the control of the cauliflower mosaic virus 35S and 19S promoters in transformed petunia tissues. *Plant Mol. Biol*. **9**, 315–324 (1987).
30. Kay, R., Chan, A., Daly, M. & McPherson, J. Duplication of CaMV-35S promoter sequences creates a strong enhancer for plant genes. *Science* **236**, 1299–1302 (1987).
31. Christensen, A. H. & Quail, P. H. Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Res*. **5**, 213–218 (1996).
32. Vain, P., Finer, K. R., Engler, D. E., Pratt, R. C. & Finer, J. J. Intron-mediated enhancement of gene expression in maize (*Zea mays* L.) and bluegrass (*Poa pratensis* L.). *Plant Cell Rep.* **15**, 489–494 (1996).
33. Stoger, E. *et al*. Cereal crops as viable production and storage systems for pharmaceutical scFv antibodies. *Plant Mol. Biol*. **42**, 583–590 (2000).
34. Artsaenko, O., Kettig, B., Fiedler, U., Conrad, U. & Düring K. Potato tubers as a biofactory for recombinant antibodies. *Mol. Breeding* **4**, 313–319 (1998).
35. Padidam, M. Chemically regulated gene expression in plants. *Curr. Opin. Plant Biol*. **6**, 169–177 (2003).
36. Padidam, M., Gore, M., Lu, D. L. & Smirnova, O. Chemical-inducible, ecdysone receptor-based gene expression system for plants. *Transgenic Res*. **12**, 101–109 (2003).
37. Cramer, C. L., Boothe, J. G. & Oishi, K. K. Transgenic plants for therapeutic proteins: linking upstream and downstream technologies. *Curr. Top. Microbiol. Immunol*. **240**, 95–118 (1999).
38. Schillberg, S., Zimmermann, S., Voss, A. & Fischer, R. Apoplastic and cytosolic expression of full-size antibodies and antibody fragments in *Nicotiana tabacum. Transgenic Res*. **8**, 255–263 (1999).
This paper compares the stability of identical scFv antibodies that are targeted to different compartments, and shows that the secretory pathway is generally much more suitable for antibody accumulation than the cytosol.
39. De Jaeger, G. *et al*. High-level accumulation of single-chain variable fragments in the cytosol of transgenic *Petunia hybrida*. *Eur. J. Biochem*. **259**, 1–10 (1998).
40. Schouten, A., Rossien, J., Bakker, J. & Schots, A. Formation of disulfide bridges by a single-chain Fv antibody in the reducing ectopic environment of the plant cytosol. *J. Biol. Chem*. **277**, 19339–19345 (2002).
41. Conrad, U. & Fiedler, U. Compartment-specific accumulation of recombinant immunoglobulins in plant cells: an essential tool for antibody production and immunomodulation of physiological functions and pathogen activity. *Plant Mol. Biol*. **38**, 101–109 (1998).
42. Plasterk, R. H. A. & Ketting, R. F. The silence of the genes. *Curr. Opin. Genet. Dev.* **10**, 562–567 (2000).
43. Anandalakshmi, R. *et al*. A viral suppressor of gene silencing in plants. *Proc. Natl Acad. Sci. USA* **95**, 13079–13084 (1998).
44. Gelvin, S. B. *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. *Microbiol. Mol. Biol. Rev*. **67**, 16–23 (2003).
45. Britt, A. B. & May, G. D. Re-engineering plant gene targeting. *Trends Plant Sci*. **8**, 90–95 (2003).
46. Veluthambi, K., Gupta, A. K. & Sharma, A. The current status of plant transformation technologies. *Curr. Sci. India* **84**, 368–380 (2003).
47. Christou, P. Transformation technology. *Trends Plant Sci*. **1**, 423–431 (1996).
48. Twyman, R. M., Stoger, E., Kohli, A. & Christou, P. in *Genetic Engineering: Principles and Practice* Vol. 24 (ed. Setlow, J. K.) 107–136 (Kluwer-Plenum, New York, 2002).
49. Martineau, B., Voelker, T. A. & Sanders, V. A. On defining T-DNA. *Plant Cell* **6**, 1032–1033 (1994).
50. Hanson, B. *et al.* A simple method to enrich an *Agrobacterium*-transformed population for plants containing only T-DNA sequences. *Plant J.* **19**, 727–734 (1999).
51. Fu, X. *et al*. Linear transgene constructs lacking vector backbone sequences generate low-copy-number transgenic plants with simple integration patterns. *Transgenic Res*. **9**, 11–19 (2000).
52. Lerouge, P., Bardor, M., Pagny, S., Gomord, V. & Faye, L. N-glycosylation of recombinant pharmaceutical glycoproteins produced in transgenic plants: towards humanisation of plant N-glycans. *Curr. Pharmaceutical Biotech.* **1**, 347–354 (2000).
53. Bardor, M. *et al*. Immunoreactivity in mammals of two typical plant glyco-epitopes: core α(1,3)-fucose and core xylose. *Glycobiology* **13**, 427–434 (2003).
54. Chargelegue, D., Vine, N. D., van Dolleweerd, C. J., Drake, P. M. & Ma, J. K. A murine monoclonal antibody produced in transgenic plants with plant-specific glycans is not immunogenic in mice. *Transgenic Res*. **9**, 187–194 (2000).
The first published paper that discusses the immunogenicity of a plant-derived glycosylated recombinant protein.
55. Warner, T. G. in *Carbohydrates in Chemistry and Biology* (eds Ernst, B., Hart, G. W. & Sanay, P.) 1043–1064 (Wiley, New York, 2000).
56. Blixt, O., Allin, K., Pereira, L., Datta, A. & Paulson, J. C. Efficient chemoenzymatic synthesis of O-linked sialyl oligosaccharides. *J. Am. Chem. Soc.* **124**, 5739–5746 (2002).
57. Bakker, H. *et al.* Galactose-extended glycans of antibodies produced by transgenic plants. *Proc. Natl Acad. Sci. USA* **98**, 2899–2904 (2001).
In this study, a transgenic tobacco line that expressed the heavy and light chains of a murine antibody was crossed with a line that expressed human β-1,4-galactosyltransferase. The progeny produced antibodies ~30% of which had partially galactosylated N-glycans, which provided a useful approach for the 'humanization' of plant glycans.
58. Raju, T. S., Briggs, J., Borge, S. M. & Jones, A. J. S. Species-specific variation in glycosylation of IgG: evidence for the species-specific sialylation and branch-specific galactosylation and importance for engineering recombinant glycoprotein therapeutics. *Glycobiology* **10**, 477–486 (2000).
In this study, cell-specific glycosylation of immunoglobulins was studied by mass spectrometry and capillary electrophoresis/laser-induced fluorescence in 13 different animal systems. The glycan patterns were found to be unique in different species, which indicated that some might be more suitable than others for the production of human therapeutic proteins. The same might apply to plant systems, which were not considered in this paper.
59. Borisjuk, N. V. *et al.* Production of recombinant proteins in plant root exudates. *Nature Biotechnol*. **17**, 466–469 (1999).
60. Komarnytsky, S., Borisjuk, N. V., Borisjuk, L. G., Alam, M. Z. & Raskin, I. Production of recombinant proteins in tobacco guttation fluid. *Plant Physiol*. **124**, 927–933 (2000).
61. Drake, P. M. W. *et al.* Rhizosecretion of a monoclonal antibody protein complex from transgenic tobacco roots. *Plant Mol. Biol.* **52**, 233–241 (2003).
62. Maliga, P. Engineering the plastid genome of higher plants. *Curr. Opin. Plant Biol*. **5**, 164–172 (2002).
63. Daniell, H., Khan, M. S. & Allison, L. Milestones in chloroplast genetic engineering: an environmentally friendly era in biotechnology. *Trends Plant Sci*. **7**, 84–91 (2002).
64. Tregoning, J. S. *et al*. Expression of tetanus toxin fragment C in tobacco chloroplasts. *Nucleic Acids Res*. **31**, 1174–1179 (2003).
65. Daniell, H., Lee, S. B., Panchal, T. & Wiebe, P. O. Expression of the native cholera B toxin subunit gene and assembly as functional oligomers in transgenic tobacco chloroplasts. *J. Mol. Biol.* **311**, 1001–1009 (2001).
66. Khan, M. S. & Maliga, P. Fluorescent antibiotic resistance marker for tracking plastid transformation in higher plants. *Nature Biotechnol*. **17**, 910–915 (1999).
67. Fischer, R., Emans, N., Schuster, F., Hellwig, S. & Drossard, J. Towards molecular farming in the future: using plant-cell-suspension cultures as bioreactors. *Biotechnol. Appl. Biochem*. **30**, 109–112 (1999).
68. Stöger, E. *et al*. Practical considerations for pharmaceutical antibody production in different crop systems. *Mol. Breeding* **9**, 149–158 (2002).
This paper considers in detail the factors that should be evaluated when choosing a crop system for the production of pharmaceutical proteins. The same scFv is expressed in many species to compare intrinsic yields, and features such as storage, distribution and biosafety are discussed, as well as economic factors.
69. Witcher, D. *et al*. Commercial production of β-glucuronidase (GUS): a model system for the production of proteins in plants. *Mol. Breeding* **4**, 301–312 (1998).
70. Hood, E. E., Woodard, S. L. & Horn, M. E. Monoclonal antibody manufacturing in transgenic plants — myths and realities. *Curr. Opin. Biotechnol*. **13**, 630–635 (2002).
71. Hood, E. E. From green plants to industrial enzymes. *Enzyme Microbial Technol*. **30**, 279–283 (2002).
72. Zeitlin, L. *et al*. A humanized monoclonal antibody produced in transgenic plants for immunoprotection of the vagina against genital herpes. *Nature Biotechnol*. **16**, 1361–1364 (1998).
73. Khoudi, H. *et al*. Production of a diagnostic monoclonal antibody in perennial alfalfa plants. *Biotechnol. Bioeng*. **64**, 135–143 (1999).
74. Perrin, Y. *et al*. Transgenic pea seeds as bioreactors for the production of a single-chain Fv fragment (scFV) antibody used in cancer diagnosis and therapy. *Mol. Breeding* **6**, 345–352 (2000).

75. De Wilde, C., Peeters, K., Jacobs, A., Peck, I. & Depicker, A. Expression of antibodies and Fab fragments in transgenic potato plants: a case study for bulk production in crop plants. *Mol. Breeding* **9**, 2871–282 (2002).
76. Schunmann, P. H. D., Coia, G. & Waterhouse, P. M. Biopharming the Simpli-RED™ HIV diagnostic reagent in barley, potato and tobacco. *Mol. Breeding* **9**, 113–121 (2002).
77. McGarvey, P. B. *et al*. Expression of the rabies virus glycoprotein in transgenic tomatoes. *Biotechnology* **13**, 1484–1487 (1995).
78. Sala, F. *et al.* Vaccine antigen production in transgenic plants: strategies, gene constructs and perspectives. *Vaccine* **21**, 803–808 (2003).
79. Commandeur, U., Twyman, R. M. & Fischer, R. The biosafety of molecular farming in plants. *AgBiotechNet* **5**, ABN 110 (2003).
80. Hare, P. D. & Chua, N.-H. Excision of selectable marker genes from transgenic plants. *Nature Biotechnol*. **20**, 575–579 (2002).
81. Zuo, J. R., Niu, Q. W., Ikeda, Y. & Chua, N. H. Marker-free transformation: increasing transformation frequency by the use of regeneration-promoting genes. *Curr. Opin. Biotechnol.* **13**, 173–180 (2002).
82. Eastham, K. & Sweet, J. *Genetically Modified Organisms (GMOs): the Significance of Gene Flow through Pollen Transfer. Environment Issue Report No. 28* (European Environment Agency, Copenhagen, 2002).
83. Kay, E., Vogel, T. M., Bertolla, F., Nalin, R. & Simonet, P. *In situ* transfer of antibiotic resistance genes from transgenic (transplastomic) tobacco plants to bacteria. *Appl. Environ. Microbiol*. **68**, 3345–3351 (2002).
84. Smalla, K. *et al.* in *Proceedings of the 6th International Symposium on the Biosafety of Genetically Modified Organisms* 146–154 (Univ. Extension Press, Univ. of Saskatchewan, Canada, 2000).
85. Smyth, S. & Phillips, P. W. B. Product differentiation alternatives: identity preservation, segregation and traceability. *AgBioForum* **5**, 30–42 (2002).
86. Schillberg, S., Fischer, R. & Emans, N. Molecular farming of recombinant antibodies in plants. *Cell. Mol. Life Sci*. **60**, 433–445 (2003).
This is a comprehensive discussion of the technical issues concerning the production of antibodies in plants, which is treated in much more detail than is possible in the present review.
87. Stoger, E., Sack, M., Fischer, R. & Christou, P. Plantibodies: applications, advantages and bottlenecks. *Curr. Opin. Biotechnol*. **13**, 161–166 (2002).
88. McCormick, A. A. *et al.* Rapid production of specific vaccines for lymphoma by expression of the tumor-derived single-chain Fv epitopes in tobacco plants. *Proc. Natl Acad. Sci. USA* **96**, 703–708 (1999).
89. Larrick, J. W., Yu, L., Naftzger, C., Jaiswal, S. & Wycoff, K. Production of secretory IgA antibodies in plants. *Biomolecular Eng*. **18**, 87–94 (2001).
This paper presents a useful summary of recent advances in the plant-based production of secretory IgAs with a discussion of purification methods and production costs.
90. Ma, J. K. *et al*. Characterization of a recombinant plant monoclonal secretory antibody and preventive immunotherapy in humans. *Nature Med*. **4**, 601–606 (1998).
91. Vaquero, C. *et al*. A carcinoembryonic antigen-specific diabody produced in tobacco. *FASEB J*. **16**, 408–410 (2002).
92. Kathuria, S. *et al.* Efficacy of plant-produced recombinant antibodies against HCG. *Human Reproduction* **17**, 2054–2061 (2002).
93. Miele, L. Plants as bioreactors for pharmaceuticals: regulatory considerations. *Trends Biotechnol*. **15**, 45–50 (1997).
94. Emlay, D. in *Plants as Factories for Protein Production* (eds Hood, E. E. & Howard, J.) 175–180 (Kluwer Academic, New York, 2002).
95. Lloyd-Evans, M. & Nair, A. in *Biopharming: the Emerging World Market of Plant-based Therapeutics*. *Theta Report No. 1214* 63–80 (PJB Medical Publications Inc., New York, 2002).
96. Stoger, E., Schillberg, S., Twyman, R. M., Fischer, R. & Christou, P. in *Methods in Molecular Biology. Antibody Engineering: Protocols and Methods* 2nd edn (ed. Lo, B. K. C.) (Humana Press Inc., New Jersey, in the press).
97. Schillberg, S., Zimmermann, S., Findlay, K. & Fischer, R. Plasma membrane display of anti-viral single chain Fv fragments confers resistance to tobacco mosaic virus. *Mol. Breeding* **6**, 317–326 (2000).
98. Sijmons, P. C. *et al.* Production of correctly processed human serum albumin in transgenic plants. *Biotechnology* **8**, 217–221 (1990).
99. Zhu, Z. *et al*. Expression of human α-interferon in plants. *Virology* **172**, 213–222 (1994).
100. Matsumoto, S., Ikura, K., Ueda, M. & Sasaki, R. Characterisation of a human glycoprotein (erythropoetin) produced in cultured tobacco cells. *Plant Mol. Biol*. **27**, 1163–1172 (1995).
101. Delaney, D. *et al*. in *Plant Biotechnology: 2002 and Beyond*. *Proceedings of the 10th IAPTC&B Congress, Orlando, Florida*. (ed. Vasil, I.) 393–394 (Kluwer Academic, Dordrecht, The Netherlands, 2002).
102. Torres, E. *et al*. Rice cell culture as an alternative production system for functional diagnostic and therapeutic antibodies. *Transgenic Res*. **8**, 441–449 (1999).
103. Terashima, M. *et al*. Production of functional human α1–antitrypsin by plant cell culture. *Appl. Microbiol. Biotechnol*. **52**, 516–523 (1999).
104. Düring, K., Hippe, S., Kreuzaler, F. & Schell, J. Synthesis and self-assembly of a functional monoclonal antibody in transgenic *Nicotiana tabacum*. *Plant Mol. Biol*. **15**, 281–293 (1990).
105. Ma, J. K.-C. *et al*. Generation and assembly of secretory antibodies in plants. *Science* **268**, 716–719 (1995).
106. Francisco, J. A. *et al*. Expression and characterization of bryodin 1 and a bryodin 1-based single-chain immunotoxin from tobacco cell culture. *Bioconjug. Chem*. **8**, 708–713 (1997).
107. Mayfield, S. P. *et al*. Expression and assembly of a fully active antibody in algae. *Proc. Natl Acad. Sci. USA* **100**, 438–442 (2003).
108. Streatfield, S. J. *et al*. Plant-based vaccines: unique advantages. *Vaccine* **19**, 2742–2748 (2001).
109. Ma, S. W. *et al.* Transgenic plants expressing autoantigens fed to mice to induce oral immune tolerance. *Nature Med*. **3**, 793–796 (1997).
110. Yu, J. & Langridge, W. H. A plant-based multicomponent vaccine protects mice from enteric diseases. *Nature Biotechnol.* **19**, 548–552 (2001).
111. Lamphear, B. J. *et al*. Delivery of subunit vaccines in maize seed. *J. Control Release* **83**, 169–180 (2002).

Acknowledgements

The authors are grateful to R. Twyman for critical assessment and help with manuscript preparation.

Online links

DATABASES

The following term in this article is linked online to:

BioCyc: http://www.biocyc.org

nos

FURTHER INFORMATION

Epicyte Pharmaceutical: http://www.epicyte.com
Large Scale Biology Corp.: http://www.lsbc.com
Medicago: http://www.medicago.com
Monsanto: http://www.monsanto.com
NeoRx: http://www.neorx.com
Pew Initiative on Food and Biotechnology: http://www.pewagbiotech.org
Phytomedics, Inc.: http://www.phytomedics.com
Prodigene: http://www.prodigene.com
SemBioSys Genetics Inc.: http://www.sembiosys.ca
Sigma Inc.: http://www.sigmaaldrich.com

Access to this interactive links box is free online.