Received: 25 April 2019 | Revised: 23 July 2019 | Accepted: 20 August 2019

DOI: 10.1002/ppp3.10073

REVIEW

Plants People Planet PPP Open Access

Therapeutic recombinant protein production in plants: Challenges and opportunities

Matthew J. B. Burnett[1] | Angela C. Burnett[2]

[1]Yale Jackson Institute for Global Affairs, New Haven, CT, USA

[2]Brookhaven National Laboratory, Upton, NY, USA

Correspondence
Angela C. Burnett, Brookhaven National Laboratory, Upton, NY, USA.
Email: aburnett@bnl.gov

Funding information
Biotechnology and Biological Sciences Research Council; Margaret Claire Ryan Fellowship Fund at the Yale Jackson Institute for Global Affairs; U.S. Department of Energy, Grant/Award Number: DE-SC0012704

Societal Impact Statement

Therapeutic protein production in plants is an area of great potential for increasing and improving the production of proteins for the treatment or prevention of disease in humans and other animals. There are a number of key benefits of this technique for scientists and society, as well as regulatory challenges that need to be overcome by policymakers. Increased public understanding of the costs and benefits of therapeutic protein production in plants will be instrumental in increasing the acceptance, and thus the medical and veterinary impact, of this approach.

Summary

Therapeutic recombinant proteins are a powerful tool for combating many diseases which have previously been hard to treat. The most utilized expression systems are Chinese Hamster Ovary cells and *Escherichia coli*, but all available expression systems have strengths and weaknesses regarding development time, cost, protein size, yield, growth conditions, posttranslational modifications and regulatory approval. The plant industry is well established and growing and harvesting crops is easy and affordable using current infrastructure. Growth conditions are generally simple: sunlight, water, and the addition of cheap, available fertilizers. There are multiple options for plant expression systems, including species, genetic constructs and protein targeting, each best suited to a particular type of therapeutic protein production. Transient expression systems provide a mechanism to rapidly transfect plants and produce therapeutic protein in a matter of weeks, rather than the months it can take for many competing expression systems, while proteins targeted to cereal seeds can be harvested, stored and potentially purified much more easily than in competing systems. Current challenges for plant expression systems include a lack of regulatory approval, environmental containment concerns and nonhuman glycosylation, which may limit the scope of the type of therapeutic proteins that can be manufactured in plants. The specific strengths of plant expression systems could facilitate the production of certain therapeutic proteins quickly and cheaply in the near future.

KEYWORDS

CRISPR/Cas9, genetic engineering, plant expression systems, protein, recombinant, therapeutic

This article has been contributed to by US Government employees and their work is in the public domain in the USA.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2019 The Authors, *Plants, People, Planet* © New Phytologist Trust

1 | INTRODUCTION

Therapeutic recombinant proteins are exogenous proteins that are expressed in a production organism and used for the treatment or prevention of disease in humans or animals. Since human insulin was first produced in *Escherichia coli* in 1982 (Kamionka, 2011; Pavlou & Reichert, 2004), therapeutic recombinant proteins have become the latest great innovation in pharmaceuticals. Since then hundreds of recombinant protein drugs have come to the market, and hundreds more are currently in development (Margolin, Chapman, Williamson, Rybicki, & Meyers, 2018; Marsian & Lomonossoff, 2016; Meyer et al., 2008; Rader, 2012; Shadid & Daniell, 2016) with the promise of treating diseases from arthritis to cancer. Unlike traditional chemically produced drugs, recombinant proteins can be very large and complex molecules with sophisticated and specific mechanisms of action. Their size and complexity make chemically synthesizing proteins incredibly difficult, so these new drugs must be produced biologically using the protein synthesis machinery found in all cells (Thomas, Deynze, & Bradford, 2002). Production using plant expression systems is both cost-effective and scalable, representing a 'major paradigm shift' for the pharmaceutical industry (Margolin et al., 2018).

The most promising therapeutic recombinant proteins are monoclonal antibodies (mAbs), originally copied from human immunoglobulin G1 (IgG1) to target epitopes with high specificity. As technology has advanced mAbs now have the potential to perform many different functions as therapeutic molecules. For instance, mAbs have the ability to stimulate the host immune system against a target cancer cell, can inhibit enzymes or inactivate other proteins and can mimic a signaling ligand or present an antigen (Dijk & Winkel, 2001). There are currently many other therapeutic recombinant proteins in production and development including hormones, growth factors, cytokines, serum proteins, enzymes, and vaccines (Margolin et al., 2018; Marsian & Lomonossoff, 2016; Rader, 2012; Shadid & Daniell, 2016).

Most therapeutic proteins are produced in either Chinese Hamster Ovary (CHO) cell cultures or *E. coli* fermentations, with a significant number also being produced in *Saccharomyces cerevisiae* and murine myeloma cells (Rader, 2008). These expression systems are the best characterized protein production platforms and each system has its own strengths and limitations. However, there are other expression systems that have not been as well utilized, which may be able to produce new therapeutic drugs or improve the production of current proteins (Table 1).

Plant cultivation technology and practice have been optimized over thousands of years to ensure high yields and low cost production for food and industry, and plant species have been domesticated to produce high biomass yields, have simple and robust growth requirements, and facilitate easier harvesting. Many of these improvements are relevant to the production of therapeutic molecules in plants, giving plant expression systems an advantage over other platforms, where much less time and money has been spent on optimization (with the possible exception of work in yeast fermentation).

While plants have not been used extensively to produce therapeutic protein products in the past, there is a history of genetically engineering plants to produce useful compounds (Vasil, 2008), and a wealth of knowledge in the scientific literature about genetic engineering in crop plants and tobacco in particular (Shinozaki et al., 1986; Zhang, Shanmugaraj, & Daniell, 2017; Zhang, Li, et al., 2017). Plants have been previously considered as expression systems for therapeutic recombinant proteins, and the concept is gathering steam again as scientists look to increase the efficiency of producing recombinant proteins (Kaiser, 2008; Tekoah et al., 2015).

Plants have many attractive characteristics as a recombinant protein production platform: cheap growth conditions, well-understood manufacturing practices, a high level of scalability, the ability to synthesize complex proteins, existing industry infrastructure, the potential for rapid production timescales, and a low risk of human pathogen contamination (Moustafa, Makzhoum, & Trémouillaux-Guiller, 2016).

In this article, we review the current status of therapeutic protein production in plants. We firstly outline the key considerations for therapeutic protein production systems, demonstrating how plants fit into the broader picture of therapeutic protein production. Next, we describe the different tools and techniques which may be used to carry out protein production in plants. We then examine the key plant species which are commonly used in this effort, and their advantages and disadvantages for therapeutic protein production. Finally, we discuss the challenges for the field of therapeutic protein production in plants and conclude by considering what the future holds for this exciting discipline.

2 | CONSIDERATIONS FOR THERAPEUTIC PROTEIN PRODUCTION SYSTEMS

There are a number of fundamental issues that must be considered when considering the most appropriate expression system to produce a therapeutic recombinant protein.

2.1 | Protein size

E. coli, or other prokaryote cells are the expression system of choice for small proteins (<30 kDa), but struggle to produce high yields of fully formed large peptides, which are more easily produced in eukaryote systems such as plants (Demain & Vaishnav, 2009).

2.2 | Folding and solubility

Correct folding of a therapeutic protein is essential for activity and complex proteins can require specific chaperone proteins to facilitate this (High, Lecomte, Russell, Abell, & Oliver, 2000; Margolin et al., 2018). Nonmammalian cells may have difficulty producing the correct folding of human proteins, especially prokaryote cells without protein processing organelles (Sahdev, Khattar, & Saini, 2008). Additionally, some expression systems (notably *E. coli*) have issues of

TABLE 1 Advantages and disadvantages of current therapeutic protein expression systems

Expression system	Advantages	Disadvantages
Bacterial	- Well characterized cell lines - Simple and cheap growth conditions - Optimized growth procedures - Scalable - Amenable to genetic engineering - Very short production timescale - Existing regulatory approval	- Nonhuman glycosylation profile - Large (>30kDa) protein mis-folding and export issues
Insect	- High protein expression levels - Scalable - Ability to produce complex eukaryotic proteins with correct folding/solubility/posttranslational modification - Existing regulatory approval	- Nonhuman glycosylation containing immunogenic sugars - Unwanted posttranslational modifications
Mammalian cell	- Correct posttranslational modifications - High yields - Many current products give precedent to regulatory bodies - Active research and industry funding - Existing regulatory approval	- Complex growth requirements raise costs - Complex cells hinder engineering and understanding - Heterologous product - Higher risk of human pathogen contamination - Unstable cell lines - Long production timescale - Difficult to scale-up
Plant	- Maximum scale-up possibility - Low growth costs - Can produce complex proteins - Low risk of contamination with human pathogens - Optimized growth procedures	- Non-human glycosylation containing immunogenic sugars - Lacks regulatory approval
Whole animal	- Massive scaling up potential - Correct posttranslational modifications - Easy harvesting - Optimized farming techniques - Stable cell lines - Low cost production - Existing regulatory approval	- Difficult and laborious to create transgenic organism - Long production timescale - Regulatory and ethical issues - Poorly characterized recombinant protein production system - Low control
Yeast/filamentous fungi	- Simple and cheap growth conditions - Fast growth to high density - Well characterized cell lines - Optimized growth procedures - Scalable - Moderately amenable to genetic engineering - Correct protein folding and processing - Short production timescale - Stable production strains - Existing regulatory approval	- Nonhuman glycosylation containing immunogenic sugars

Note: Information sourced from Collares, Bongalhardo, Deschamps, & Moreira, 2005; Demain & Vaishnav, 2009; Ghaderi et al., 2012; Gomes et al., 2019; Lagassé et al., 2017; Ma et al., 2003; Verma et al., 1998; Walsh, 2010.

insoluble protein accumulation when the product is overexpressed (Verma, Boleti, & George, 1998).

2.3 | Posttranslational modification

After translation, many proteins are modified and these modifications may include the formation of covalent bonds, as in the case of disulphide bridges, or the addition of carbohydrate molecules in a process known as glycosylation (Box 1). Most of these posttranslational modification (PTM) mechanisms are conserved across eukaryotes and prokaryotes, but glycosylation mechanisms can differ even between species. Many secretory human proteins are glycosylated which can be essential for protein function affecting serum half-life, immunogenicity, effector function, and solubility (Lim et al., 2010; Sethuraman & Stadheim, 2006). This raises a problem for expression systems based on nonhuman cells, therefore glycosylation is a major concern for every expression system (Box 1, Table 1). The capacity of plants to carry out glycosylation is an advantage over prokaryotic expression systems; and even in insect and yeast cells, glycosylation capacity is limited (Marsian & Lomonossoff, 2016). Glycoengineering in all of the available systems aims to increase the production of human-like glycosylation

Box 1. Nonhuman glycosylation profiles

Each expression system faces its own glycosylation challenges. *Escherichia coli* does not possess any native glycosylation machinery and when engineered to express a *Campylobacter jejuni* glycosylation system can only glycosylate fully folded proteins (Kowarik et al., 2006), although this can be overcome in some cases using chemical modification (e.g., PEGylation) (DeFrees et al., 2006). Yeast expression systems can glycosylate, but glycan molecules have a much higher proportion of mannose residues than human glycans and often lack fucose and terminal sialic acid residues, reducing the half-life in patients (Ghaderi, Zhang, Hurtado-Ziola, & Varki, 2012; Walsh, 2010). Insect cells add paucimannosidic glycans, which are not found in humans. Plants exhibit a range of different glycosylation mechanisms which lack certain sugars, including terminal sialic acid residues, and often include β1-2xylose and α1-3fucose residues, which elicit an immune response when introduced intravenously (Gomord, Chamberlain, Jefferis, & Faye, 2005; Gomord et al., 2010; Walsh, 2010). Even mammalian (nonhuman) expression systems do not exactly mimic human glycosylation, adding Gala1-3Gal (alpha-Gal) and N-glycolylneuraminic acid (Neu5Gc) residues, which cause rapid clearance of the protein from the bloodstream (Varki, 2009). Homogeneity is a desirable characteristic of any therapeutic molecule, and consistent glycosylation profiles are a challenge for mammalian cell expression systems in particular (Sethuraman & Stadheim, 2006).

profiles in recombinant proteins, and the success of this work may determine the success of individual expression systems in the future (Montero-Morales & Steinkellner, 2018; Sethuraman & Stadheim, 2006).

2.4 | Safety

As mentioned above, nonhuman PTMs can cause an immune response against the therapeutic protein, and some expression systems have a risk of introducing other contaminants into the drug. Mammalian expression systems have a higher risk of transferring pathogens (e.g., viruses or prions) to the patient (Lico, Santi, Twyman, Pezzotti, & Avesani, 2012). Bacterial expression systems risk introducing toxins such as O-antigen (Fischer & Emans, 2000). Plant systems generally avoid both of these pitfalls. These risks must be addressed with purification procedures, adding to the cost of downstream processing.

2.5 | Genetic engineering

All of the expression systems require the use of transgenic organisms/cell lines, so the ease and stability of performing genetic engineering is particularly relevant. Expression systems that are well characterized and have many genetic tools, such as expression vectors and strong promoters optimized for use in that specific system, will have an advantage. Producing a transgenic *E. coli* is much easier (Verma et al., 1998) than producing a transgenic goat because of the complexity of the goat's genome and because genetic manipulation is well understood in *E. coli*. CHO cells, *S. cerevisiae*, and *E. coli* are the best understood and therefore the most used expression systems. Using a well-characterized system reduces development time and increases the predictability of the production process. Even CHO cells, a well characterized mammalian cell type, rely on essentially random integration of expression cassettes (Barnes, Bentley, & Dickson, 2003; Manivasakam, Aubrecht, Sidhom, & Schiestl, 2001), and in these less controlled genetic engineering approaches detailed screening is the key to creating productive strains. Another consideration is the genetic stability of an expression system (Barnes et al., 2003), which determines how long the system will continue to produce the target protein at the original level and specificity. Plant expression systems are relatively easy to manipulate genetically, and transgenes are generally more stable than in bacterial systems.

2.6 | Yield

The maximum yield of each system is a major consideration. It is obviously beneficial to get the highest yield of correctly folded and posttranslationally modified protein from an expression system, but this is particularly important with regard to downstream processing, which becomes significantly more expensive when purifying protein from a more dilute mixture. The type of cell used to produce the therapeutic protein will also affect the purification procedures used in downstream processing, affecting the overall yield and cost of processing (Kozlowski & Swann, 2006).

2.7 | Growth conditions and rate

The growth rate will significantly affect the productivity of each system as production is usually run in a batch process. A faster growth rate will allow more batches over a set time. The specific growth requirements also affect the cost of a process, some cell types, for example, yeast or bacteria, can be grown to a high concentration on a cheap, simple media, while others, such as mammalian cells, require very complex and expensive media for optimum growth.

With great variation between expression systems, and the large number of different therapeutic recombinant proteins, it is unlikely that there is a 'one system fits all' solution to producing affordable protein drugs. In the same way that smaller proteins are currently produced in *E. coli* and larger proteins requiring human-like posttranslational modifications are produced in CHO cells, different systems are likely to prove to be the most effective expression systems for different proteins. Plants may prove to be the ideal in-between system, able to produce larger therapeutic proteins than bacteria,

while being more scalable and cost-effective than mammalian cell systems, as well as reducing the risk of pathogens and toxic contaminants compared to both systems.

3 | TOOLS AND TECHNIQUES FOR PROTEIN PRODUCTION IN PLANTS

There is a wide range of options available when choosing a plant expression system, ranging from the choice of expression vector and promoter to the type of plant that will be used. These options can generate huge differences in yield, protein storage capacity, ease of harvest, and posttranslational modification and must be chosen carefully to suit the requirements for the production of each specific recombinant protein.

3.1 | Expression types

Optimal yield of recombinant protein relies on a controlled, high level of transcription, translation, correct folding, targeting, and protein stability (Ma, Drake, & Christou, 2003). The keys to high levels of transcription are the regulatory genetic elements, the most important of which in plants are the promoter and the polyadenylation site.

3.1.1 | Nuclear expression

The basic expression system incorporating transgenes into the nuclear genome of a plant, nuclear expression is the conventional method of genetically engineering plants (Figure 1). Nuclear expression involves transcription in the nucleus and translation in the cytoplasm. It involves the expression of a foreign antigen from the nuclear genome, introduced into the plant using either *Agrobacterium tumefaciens*-mediated transformation or biolistic gene gun-mediated transformation; signal peptides are used to target proteins for secretion or organellar storage (Shadid & Daniell, 2016). This is the simplest and most widely used method of genetically modifying crops. Disadvantages of this system include gene silencing, risk of transgene contamination through reproductive tissues, and low expression levels (Shadid & Daniell, 2016). A further disadvantage is the need for time-consuming genetic manipulation procedures requiring backcrosses and plant breeding for the expression of multiple genes and random integration of genes which can adversely affect expression. This could be overcome by novel gene editing techniques such as CRISPR-Cas9 (Gomes, Oliveira, Vieira, & Duque, 2019; Jaganathan, Ramasamy, Sellamuthu, Jayabalan, & Venkataraman, 2018; Miki, Zhang, Zeng, Feng, & Zhu,

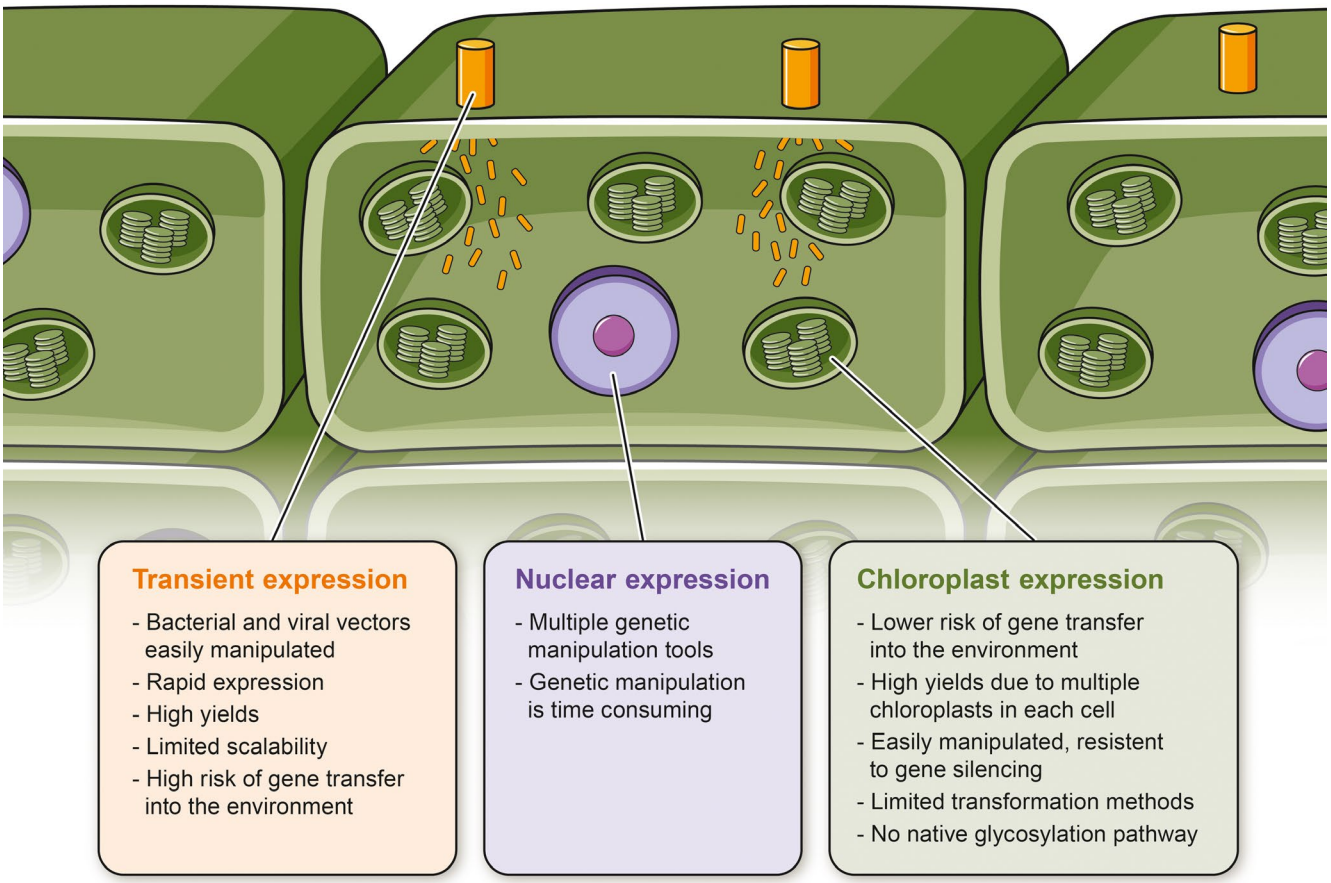

FIGURE 1 Simplified plant cell diagram showing localization and features of transient (yellow), nuclear (purple) and chloroplast (green) expression systems. For clarity, size of transient vectors and chloroplasts have been exaggerated, and additional organelles have been omitted from the diagram

2018). This system, a component of immunity in bacteria, uses clustered regularly interspaced short palindromic repeats (CRISPR) alongside the prokaryotic-traceable RNA-guided nuclease Cas9, to precisely edit the genome, and has been applied in both prokaryotes and eukaryotes as a mechanism of genome editing. CRISPR/Cas9 requires co-transformation of two vectors, which give rise to a crRNA and a tracRNA; these form a two-RNA structure and integrate to form one transcript, the sgRNA, which guides the Cas9 endonucleases to the target DNA sequences (Wang et al., 2018). CRISPR/Cas9 is highly efficient and highly robust, for example when compared to zinc finger nucleases and transcription activator-like effector nucleases, and is site-specific. A recent study in cotton showed no off-target editing and reported genome editing with an efficiency of 66.7%–100% at each of multiple sites (Wang et al., 2018); off-target mutations seem to occur more frequently in human cells than in plant cells. The most popular promoter for use in dicots is the CaMV 35S from the cauliflower mosaic virus (Ma et al., 2003), a strong constitutive promoter which can be boosted by duplicating the enhancer region (Kay, Chan, Daly, & McPherson, 1987). Alternative promoters such as the maize ubiquitin-1 promoter are used effectively in monocots (Ma et al., 2003; Twyman, Stoger, Schillberg, Christou, & Fischer, 2003). A variety of polyadenylation sequences can be used, the *Agrobacterium tumefaciens nos* gene, the pea *ssu* gene and the cauliflower mosaic virus 35S transcript being popular examples. Polyadenylation is one of the major factors determining expression levels, and is important for export of mRNA from the nucleus and subsequent translation, as well as being a key element of mRNA stability (Ma et al., 2003). Strong, constitutive promoters may give a high overall protein yield, but more nuanced approaches are being explored, as documented in the literature (Ma et al., 2003; Twyman et al., 2003). Tissue-specific promoters, such as those expressed in cereal seeds, target the protein production to certain tissues allowing easier harvesting of the product and avoiding toxicity in the parent plant which may inhibit growth (Twyman et al., 2003). In fact, with the discovery of a nectary promoter, work has been done to express proteins in the nectar of a flower, which can be harvested by bees and concentrated into honey (Breithaupt, 1999). Honey has the multiple advantages of concentrating the protein and being made up of almost exclusively sugar, greatly easing the purification process. Honey also has natural preservative properties, increasing the shelf-life of the protein (Breithaupt, 1999). Inducible promoters have also been used to initiate protein production just before, or after harvest, again, to avoid the growth limiting effects of recombinant protein over-expression (Twyman et al., 2003).

3.1.2 | Chloroplast expression

Chloroplast expression involves the introduction of a transgene into the chloroplast genome using a particle gun. Transforming a recombinant gene into the chloroplast genome has a number of advantages over nuclear transformation (Figure 1). The chloroplast genome is more easily manipulated—if the chloroplast genome has been sequenced, a transgene cassette can be created to insert foreign genes into a spacer region between functional chloroplast genes, using two known flanking sequences in the chloroplast genome, via homologous recombination (Daniell, Lin, Yu, & Chang, 2016; Daniell, Streatfield, Streatfield, & Wycoff, 2001). This precise targeting avoids placing the gene into a part of the genome which is poorly transcribed, ensuring a high level of expression. Additionally, gene silencing has not been documented using this method. Transformation into the chloroplast genome is more difficult than transformation into the nuclear genome due to the double membrane barrier found around the chloroplast and the lack of any virus known to infect the chloroplast. However, effective transformation has been achieved using the gene gun method—bombarding young plant tissue with gold or tungsten particles coated with DNA (Verma, Samson, Koya, & Daniell, 2008). Since there are thousands of copies of the chloroplast genome in each leaf cell, very high yields (over 70% of the total soluble protein in plant leaves) have been achieved using chloroplast expression (Daniell et al., 2016) as the method allows a high gene copy number per cell (Ma et al., 2003; Shadid & Daniell, 2016). Chloroplast expression has the added benefit of reducing the risk of genes leaching into the environment as chloroplast genes are maternally inherited in most crop plants, and expression in the chloroplasts allows harvest before the appearance of any reproductive structures ensuring "total biological containment of transgenes" (Verma et al., 2008). Glycosylation does not occur in chloroplasts, which allows the production of therapeutic proteins completely free of glycosylation (Verma et al., 2008). This removes a source of immunogenicity but also limits the ability to produce some therapeutic proteins such as antibodies which require glycosylation to function. Conversely the lack of a glycosylation pathway provides a glycoengineering opportunity, with a "clean slate" to engineer a custom glycosylation mechanism in chloroplasts without the need to alter or interfere with host glycosylation pathways which may be essential for cell viability. The current chloroplast expression system is best suited to proteins which do not require significant posttranslational modification and a number of vaccines and human proteins have been produced using this method including cholera toxin B (Daniell, Lee, Lee, Panchal, & Wiebe, 2001), tetanus toxin fragment c (Tregoning, 2003), anthrax protective antigen (Watson, Koya, Leppla, & Daniell, 2004), human serum albumin (Fernández-San Millán, Mingo-Castel, Miller, & Daniell, 2003), and human somatotropin (Staub et al., 2000); further viral and bacterial antigens that have been expressed in the chloroplast genome are summarized by Shadid and Daniell (2016).

Transgenes are commonly integrated between the *trnI-trnA* genes in the *rrn* operon, as this is a transcriptionally active region that offers very high levels of gene expression. Commonly used sequences in plasmid gene vectors include the bacteriophage T7 gene 10 as a 5' untranslated region to enhance ribosome binding, the use of a 3' untranslated region to ensure transcript stability, and the use of a chloroplast promoter such as *psbA* (Daniell & Jin, 2015).

3.1.3 | Transient expression

Transient expression (Figure 1) allows the rapid production of recombinant proteins, drastically reducing the development time of the expression system. This can be used to test genetic constructs and for rapid sampling of recombinant proteins for functional analysis (Twyman et al., 2003). Transient expression also has the potential to be used for the production of large amounts of protein as a mainstream production platform, but ultimately has limited scaling up potential compared with transgenic plants (Vaquero et al., 2002). The rapidity of the system nevertheless provides the potential for a rapid response, for example, in response to a pandemic, since the need for full transformation is eliminated (Marsian & Lomonossoff, 2016). For example, purified end product of an influenza vaccine was produced just three weeks after the sequence was received by Medicago, a company specialized in plant-based transient expression systems (D'Aoust et al., 2010). There are two established transient expression methods. In the first of these, plant viruses such at the tobacco mosaic virus are used to introduce the transgene into an infected plant (Shih & Doran, 2009), but there is a risk of the viral vector infecting plants in the ecosystem. The second method involves *Agrobacterium* mediated transient gene expression introducing T-DNA into plant cells for high level and high efficiency expression (Kapila, Rycke, Montagu, & Angenon, 1997). Transient expression using this system is much more efficient than that of integrated genes, reported to be at least 1,000 fold higher (Janssen & Gardner, 1990) with yields reported at up to 1.5 g of antibody per kg of leaf (fresh weight) (Vézina et al., 2009). *Agrobacterium*-mediated transient expression (Agroinfiltration), has the benefit of reaching a very high percentage of cells in a treated tissue (Obembe, Popoola, Leelavathi, & Reddy, 2011), whereas viral infection can be limited to the outer layer of cells. Magnifection is a combination of the two transient expression methods developed by Icon Genetics, using *Agrobacterium* to deliver viral vectors (Obembe et al., 2011). Magnifection not only increases infectivity and therefore coverage of the plant, but also increases yield and allows the co-expression of multiple proteins required for the assembly of hetero-oligomeric proteins (Giritch et al., 2006). An example of the utility of transient expression in plants using *Agrobacterium*-mediated transformation is the expression of viral coat proteins, which assemble into virus-like particles (VLPs). VLPs do not contain infectious genomic material, so they are considered safe, yet they are similar enough to virus particles to successfully elicit an immune response (Marsian & Lomonossoff, 2016). The safety of these particles is a major advantage over traditional vaccine production—relying on attenuated or inactivated pathogens carries an inherent risk of incomplete attenuation or inactivation. Hepatitis B VLPs were one of the first particles to be produced using transient expression in plants, and a wide variety of VLPs have since been produced in plants, giving a positive immune response in animal models. For example, bovine papillomavirus VLPs expressed in *Nicotiana benthamiana* successfully elicited a positive immune response in rabbits (Love et al., 2012). In addition to immunity, VLPs could also be used for drug delivery (Marsian & Lomonossoff, 2016). eVLPS (RNA-free, empty VLPs) are being developed for various applications including cell-specific drug targeting (Wen et al., 2012). Successful VLP production may partly depend on ensuring acid- and thermostability, for both function and storage purposes; recent work has shown that site-directed mutagenesis used to introduce amino acid substitutions increasing acid–and thermostability increased the stability and yield of VLPs engineered in *Nicotiana benthamiana* leaves (Veerapen, Zyl, Rybickia, & Meyersa, 2018).

3.1.4 | Suspension cells

Suspension cell cultures have the same advantages of sterility, containment, and well-defined downstream processing procedures which other cell culture expression systems possess, but lose many of the aspects of plant expression systems that make them attractive including the huge scaling up potential (Twyman et al., 2003). The ability to use low cost defined growth media is an advantage over mammalian cell culture, but therapeutic protein production using plant cells in suspension offers few advantages over a yeast or insect expression system.

3.2 | Production species

3.2.1 | Tobacco

The molecular biology workhorse of the plant world, tobacco is the most widely used species for the production of recombinant proteins in the laboratory (Ma et al., 2003). Benefits of using tobacco include a high biomass yield of "more than 100,000 kg per hectare for close-cropped tobacco" (Ma et al., 2003), and rapid scale up potential due to a huge seed production capacity. Protein storage in the leaves is not particularly stable and the product is vulnerable to degradation, so the leaves must be frozen or dried for storage or the protein extracted soon after expression. Tobacco tissues usually contain phenols and toxic alkaloids which have implications for downstream processing.

3.2.2 | Cereals

Cereal seeds are excellent protein storage devices equipped with protein storage vesicles and a dry intercellular environment, reducing protease activity and the rate of nonenzymatic hydrolysis. Maize has the highest biomass yield among food crops (Obembe et al., 2011) and has already been used in the production of avidin (Hood et al., 1997), bovine trypsin and recombinant antibodies to name a few (Ma et al., 2003). Dry cereal seeds such as those from rice and wheat have the advantage of high protein stability, allowing storage at room temperature for a matter of months without significant loss of activity (Stöger et al., 2000); additionally, rice is self-fertilizing, reducing the risk of transgenes being transferred to other plants (Rybicki, 2010). Food crops also present the opportunity to administer oral vaccines produced in the crop by feeding them to patients with minimal processing (Margolin et al., 2018). Coupled with the

stability of proteins in seeds, this presents an extremely attractive opportunity to reduce the cost and distribution issues faced by conventional vaccines (Stöger et al., 2000). However, with strict regulatory requirements, it is unlikely that an edible plant vaccine could be used in humans without a level of processing and formulation to homogenize the product and make sure the correct dose and potency was reproducible in all products (Rybicki, 2010). The concept of producing vaccines in food crops has lost favor in recent years after two incidents in the USA where transgenic plant material contaminated wild-type food crops. These incidents have resulted in a tightening of regulations and a reduced interest from drug companies to pursue the production of vaccines in edible crops (Rybicki, 2010) although edible vaccines against *E. coli*, produced by potato and maize, have reached phase I clinical trials (Shadid & Daniell, 2016).

3.2.3 | Legumes

Therapeutic protein production has been documented in legumes such as soybean, pea, and alfalfa. Legumes have the advantage of fixing atmospheric nitrogen, removing the nitrogen requirement in their fertilizer, and therefore reducing cultivation cost. However, these plants do have lower leaf biomass than tobacco (Ma et al., 2003). Grain legumes such as peas have high protein content in their seeds, and are being developed as expression systems (Perrin et al., 2000).

3.2.4 | Fruits and vegetables

A number of fruit and vegetable crops have been used to produce therapeutic recombinant proteins, including lettuce, tomato, and most frequently, potato. Like for cereals, a great advantage of these systems is that the protein could be delivered orally with minimal processing, although as mentioned previously guaranteeing the dose and quality is a challenge (Daniell, Kulis, & Herzog, 2019; Ma et al., 2003; Marsian & Lomonossoff, 2016; Rybicki, 2010).

4 | CHALLENGES FACED BY PLANT EXPRESSION SYSTEMS

As attractive as plants may seem as therapeutic protein expression systems, there are a number of challenges that must be overcome before they can be widely adopted.

4.1 | Environmental contamination

Perhaps the biggest challenge facing protein expression in plants are the concerns around genetically modified (GM) crops. Major concerns include the spread of recombinant genes through seed dispersal, pollen dispersal, viral transfer or horizontal transfer; therapeutic proteins getting into the food supply of humans or animals; and adverse effects on organisms in the environment (Ma et al., 2003; Obembe et al., 2011). In recent years, USDA legislation has reacted to incidents of transgenic plants being found in food crops (Kaiser, 2008; Ma et al., 2003; Rybicki, 2010). There are a number of strategies that can be used to ease these concerns including geographical containment, using different planting seasons than those of local food crops, the use of male sterility in GM plant strains, using the chloroplast expression system (Lau & Sun, 2009), the use of inducible promoters, producing easily identified plant varieties (e.g., white tomatoes) (Ma et al., 2003), using self-pollinating species, producing nongerminating seeds (Obembe et al., 2011), and producing inactive fusion proteins that are activated by postpurification processing (Daniell, Streatfield, et al., 2001). Growing crops inside appropriately managed greenhouses, hydroponic growth rooms or using cell suspension cultures can provide an effective and economical means of containing GM plant material (Ma et al., 2003; Obembe et al., 2011; Su et al., 2015).

4.2 | Regulatory approval

As promising as this technology may be, drug companies are unwilling to risk the huge sums of money required to get a new product approved by the large drug approval administrations if there is already a proven alternative expression system with regulatory approval (Rybicki, 2010). This economic constraint has a stagnating effect on the pharmaceutical industry, limiting the scale of progress and the development of new drug production technologies. Unfortunately this situation is unavoidable because of the high level of confidence that is needed in any therapeutic molecule to be used in humans. The production of animal vaccines in plants is making faster progress, as there are fewer regulatory hurdles (Rybicki, 2010); this could provide a proof-of-concept for the production of human vaccines in plants, demonstrating the value of the expression system to produce effective therapeutic proteins cost effectively. A large advantage which plant expression systems have over conventional therapeutic protein production platforms is the ability to produce protein rapidly, going from gene sequence to grams of protein in under a month using transient expression techniques (Rybicki, 2010). This is preferable to the current influenza vaccine production system using eggs which 'does not provide sufficient capacity and adequate speed to satisfy global needs to combat newly emerging strains, seasonal or potentially pandemic' (Shoji et al., 2011). This provides a significant advantage over conventional methods of responding to rapidly emerging disease strains, as was shown in 2014 when an Ebola treatment was produced at short notice in *Nicotiana benthamiana* using a transient expression system (Gomes et al., 2019). This is an opportunity for plant expression systems to excel, producing vaccines quickly in response to emerging threats such as rapidly mutating diseases or bioterror threats. In the case of the Ebola treatment, full regulatory approval was sidestepped under compassionate protocols (Gomes et al., 2019). The first plant-produced therapeutic protein to win full regulatory approval for human use was taliglucerase alpha produced in carrot cell culture (Tekoah et al., 2015). The molecule was already approved from mammalian cell

culture, so it was easier to transfer approval to a new production system than to bring an entirely new product through the regulatory process (Gomes et al., 2019). These advances will undoubtedly make it easier for further drugs to be licenced in future and pharmaceutical companies should now be more likely to consider plant expression systems (Davies, 2010).

4.3 | Protein stability

The stability of expressed proteins is a concern which has significant bearing on the overall viability of the expression system. The solutions to unstable protein breakdown are dependent on the individual recombinant protein being expressed, but could include the following: the creation of fusion proteins with a stabilizing peptide co-expressed with the therapeutic protein (this method can also facilitate downstream processing with the use of affinity tags); protein targeting to seeds, oil bodies or protein storage vacuoles; freeze-drying plant material in order to preserve expressed proteins; and for proteins that do not require glycosylation the chloroplast expression system is ideal for maximizing protein yield, stability and accumulation (Daniell et al., 2019; Obembe et al., 2011).

4.4 | Posttranslational modifications

The plant proteome is highly plastic, facilitating extensive engineering: the simultaneous co-expression of many proteins enables complex protein production pathways to be established, allowing the possibility of complex glycosylation engineering (Margolin et al., 2018). Whilst plants have a similar glycosylation mechanism to humans, there are differences in terms of N-glycan composition—notably the addition of α1-3fucose and β1-2xylose and the absence of α1-6fucose, glucose and sialic acid residues (Obembe et al., 2011). These differences can have drastic effects on the distribution, half-life in serum, activity, and immunogenicity of therapeutic proteins (Twyman et al., 2003). While safety concerns may be unwarranted (Ma et al., 2003), there is no doubt that consistent human-like N-glycosylation is a vital goal in the production of some therapeutic proteins such as monoclonal antibodies (Raju, Briggs, Borge, & Jones, 2000). However, there are therapeutic proteins which may not require such specific posttranslational modification, and these proteins may be better suited to production in plants. There are several strategies proposed to overcome the problem of nonhuman N-glycosylation: in vitro modification using purified human β1-4 galactosyltransferase and sialyltransferase enzymes (Blixt, Allin, Pereira, Datta, & Paulson, 2002), knock-out/knock-down of the native plant fucosyltransferase and xylyltransferase enzymes (Twyman et al., 2003), and expressing human β1-4 galactosyltransferase in the transgenic plant (Bakker et al., 2001). Recombinant viral structural proteins may be readily produced in plants, but viral glycoproteins pose a similar challenge to mammalian glycoproteins (Margolin et al., 2018). The issue of glycosylation, whilst challenging, is not insurmountable: plant-derived influenza haemagglutinin, the only viral glycoprotein to have been tested in humans, has successfully been engineered with glycans at all possible sites and is anticipated to have FDA approval by 2020 (Le Mauff et al., 2015; Margolin et al., 2018; Ward et al., 2014); a suite of viral glycoprotein vaccine candidates against a range of diseases—including influenza, HIV, and Ebola—have been expressed in plants, summarized by Margolin et al. (2018). Finally, chloroplast expression provides a 'blank slate' for in vitro or in vivo glycoengineering without interfering with the native glycosylation mechanism. Although the ability of chloroplasts to add posttranslational modifications is not fully understood, they have been shown to have the capabilities for phosphorylation, lipidation and forming disulphide bonds (Zhang, Shanmugaraj, & Daniell, 2017; Zhang, Li, et al., 2017).

5 | PERSPECTIVES AND FUTURE DIRECTIONS

The potential market for therapeutic proteins is huge, with products ranging from antibodies to hormones and enzymes to vaccines. Each type of recombinant protein has its own production challenges and these will inevitably match up with the strengths of the different expression systems available.

The relatively short time it takes to go from sequence to producing grams of protein, using high yield transient expression systems such as Magnifection is a major advantage plants have over other expression systems. This strength lends itself to the production of vaccines to treat emerging or rapidly mutating diseases such as influenza or bioterror threats. There is also the potential for small production runs using this technology, for orphan diseases with a small number of patients, or perhaps even personalized treatments. The rapid production combined with the ability to grow transgenic plants in low cost greenhouses could greatly reduce the otherwise high cost of protein drugs for rare diseases.

As the therapeutic protein market matures, patents will expire, allowing the production of "biosimilars"—copies of the original, licensed protein produced off patent (Davies, 2010). Plant expression systems, for example a high yield chloroplast expression system, could allow the production of these proven drugs on a much larger scale and at a lower cost, grown in greenhouses or perhaps in the field (with the appropriate containment strategies in place). With the current state of glycoengineering in plants these therapeutic proteins could not require essential, human-like N-glycosylation as this is not yet available in plants (Strasser, 2016). But with progress in engineering, the glycosylation pathway, and in vitro glycosylation procedures, N-glcosylated therapeutic proteins produced in plants could be a possibility in the near future.

One of the largest barriers to widespread acceptance of plant expression systems is the lack of regulatory approval, although there are plant produced recombinant protein products on the market most are either diagnostic, veterinary or classed as medical devices, which are not required to meet the high standards of drugs for human use (Lico et al., 2012). The difficulty and cost of gaining this approval currently outweighs the benefits of using plants to produce therapeutic proteins. One supposed benefit of

plant expression systems is low cost and high scalability. While it is true that plants have the potential to produce more protein more cheaply than mammalian cell culture, for example, this only has a limited impact on the overall cost of producing a therapeutic protein drug. The major part of the cost is in purification of the product, which would essentially be the same in the cell extract of a plant or mammalian cell. If protein harvest and purification could be done at a lower cost in plants, most likely through targeting the expression to certain storage bodies such as seeds, which have a lower volume of water, or nectar which has few other contaminants from which to extract the protein, the economic benefit of using a plant expression system would be much greater. Alternatively, if purification can be sidestepped entirely such as in the example of coagulation factor IX in lettuce leaves for the treatment of hemophilia B, plant expression systems become hugely attractive (Su et al., 2015).

Plants may also be considered safer than many other expression systems, since they do not constitutively produce endotoxins, or naturally support the growth of viruses or prions with the potential for infecting humans (Moustafa et al., 2016).

As the understanding of recombinant protein expression systems increases and their limitations are fully understood, companies will be able to make informed choices on the ideal expression systems available to produce a specific therapeutic protein. Plant expression systems will no doubt fit into this landscape, but how much they are utilized relies on how effectively the challenges can be overcome.

ACKNOWLEDGEMENTS

M.J.B.B. was supported by a Biotechnology and Biological Sciences Research Council Master's degree scholarship at the University of Sheffield and by the Margaret Claire Ryan Fellowship Fund at the Yale Jackson Institute for Global Affairs. A.C.B. was supported by the United States Department of Energy contract No. DE-SC0012704 to Brookhaven National Laboratory.

AUTHOR CONTRIBUTIONS

M.J.B.B. carried out the literature review, drafted the manuscript, and prepared the figure, box and table. M.J.B.B. and A.C.B. developed the manuscript. A.C.B. prepared the manuscript for publication.

ORCID

Angela C. Burnett https://orcid.org/0000-0002-2678-9842

REFERENCES

Bakker, H., Bardor, M., Molthoff, J. W., Gomord, V., Elbers, I., Stevens, L. H., … Bosch, D. (2001). Galactose-extended glycans of antibodies produced by transgenic plants. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(5), 2899–2904. https://doi.org/10.1073/pnas.031419998

Barnes, L. M., Bentley, C. M., & Dickson, A. J. (2003). Stability of protein production from recombinant mammalian cells. *Biotechnology and Bioengineering*, *81*(6), 631–639. https://doi.org/10.1002/bit.10517

Blixt, O., Allin, K., Pereira, L., Datta, A., & Paulson, J. C. (2002). Efficient chemoenzymatic synthesis of O-linked sialyl oligosaccharides. *Journal of the American Chemical Society*, *124*(20), 5739–5746. https://doi.org/10.1021/ja017881+

Breithaupt, H. (1999). A spoon full of honey. *Nature Biotechnology*, *17*(9), 838–838. https://doi.org/10.1038/12795

Collares, T., Bongalhardo, D. C., Deschamps, J. C., & Moreira, H. L. M. (2005). Transgenic animals : The melding of molecular biology and animal reproduction. *Animal Reproduction*, *2*, 11–27.

Daniell, H., & Jin, S. (2015). The engineered chlorpoplast genome just got smarter. *Trends in Plant Science*, *20*(10), 622–640.

Daniell, H., Kulis, M., & Herzog, R. (2019). Plant cell-made protein antigents for induction of Oral tolerance. *Biotechnology Advances*, https://doi.org/10.1016/j.biotechadv.2019.06.012

Daniell, H., Lee, S. B., Panchal, T., & Wiebe, P. O. (2001). Expression of the native cholera toxin B subunit gene and assembly as functional oligomers in transgenic tobacco chloroplasts. *Journal of Molecular Biology*, *311*(5), 1001–1009. https://doi.org/10.1006/jmbi.2001.4921

Daniell, H., Lin, C., Yu, M., & Chang, W. (2016). Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. *Genome Biology.*, *17*, 134. https://doi.org/10.1186/s13059-016-1004-2

Daniell, H., Streatfield, S. J., & Wycoff, K. (2001). Medical molecular farming: Production of antibodies, biopharmaceuticals and edible vaccines in plants. *Trends in Plant Science*, *6*(5), 219–226. https://doi.org/10.1016/S1360-1385(01)01922-7

D'Aoust, M., Couture, M. M., Charland, N., Trépanier, S., Landry, N., Ors, F., & Vézina, L. (2010). The production of hemagglutinin-based virus-like particles in plants: A rapid, efficient and safe response to pandemic influenza. *Plant Biotechnology Journal*, *8*, 607–619. https://doi.org/10.1111/j.1467-7652.2009.00496.x

Davies, H. M. (2010). Commercialization of whole-plant systems for biomanufacturing of protein products: Evolution and prospects. *Plant Biotechnology Journal*, *8*(8), 845–861.

DeFrees, S., Wang, Z.-G., Xing, R., Scott, A. E., Wang, J., Zopf, D., … Clausen, H. (2006). GlycoPEGylation of recombinant therapeutic proteins produced in *Escherichia coli*. *Glycobiology*, *16*(9), 833–843. https://doi.org/10.1093/glycob/cwl004

Demain, A. L., & Vaishnav, P. (2009). Production of recombinant proteins by microbes and higher organisms. *Biotechnology Advances*, *27*(3), 297–306. https://doi.org/10.1016/j.biotechadv.2009.01.008

Fernández-San Millán, A., Mingo-Castel, A., Miller, M., & Daniell, H. (2003). A chloroplast transgenic approach to hyper-express and purify human serum albumin, a protein highly susceptible to proteolytic degradation. *Plant Biotechnology Journal*, *1*(2), 71–79. https://doi.org/10.1046/j.1467-7652.2003.00008.x

Fischer, R., & Emans, N. (2000). Molecular farming of pharmaceutical proteins. *Transgenic Research*, *9*(4–5), 279–299; Discussion 277.

Ghaderi, D., Zhang, M., Hurtado-Ziola, N., & Varki, A. (2012). Production platforms for biotherapeutic glycoproteins. Occurrence, impact, and challenges of non-human sialylation. *Biotechnology and Genetic Engineering Reviews*, *28*(1), 147–175. https://doi.org/10.5661/bger-28-147

Giritch, A., Marillonnet, S., Engler, C., van Eldik, G., Botterman, J., Klimyuk, V., & Gleba, Y. (2006). Rapid high-yield expression of full-size IgG antibodies in plants coinfected with noncompeting viral vectors. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(40), 14701–14706. https://doi.org/10.1073/pnas.0606631103

Gomes, C., Oliveira, F., Vieira, S. I., & Duque, A. S. (2019). Prospects for the production of recombinant therapeutic proteins and peptides

in plants: special focus on angiotensin I-converting enzyme inhibitory (ACEI) peptides. In F. Jamal (Ed) *Genetic EngineeriSng*. London: IntechOpen. https://doi.org/10.5772/intechopen.84419

Gomord, V., Chamberlain, P., Jefferis, R., & Faye, L. (2005). Biopharmaceutical production in plants: Problems, solutions and opportunities. *Trends in Biotechnology*, *23*(11), 559–565. https://doi.org/10.1016/j.tibtech.2005.09.003

Gomord, V., Fitchette, A.-C., Menu-Bouaouiche, L., Saint-Jore-Dupas, C., Plasson, C., Michaud, D., & Faye, L. (2010). Plant-specific glycosylation patterns in the context of therapeutic protein production. *Plant Biotechnology Journal*, *8*(5), 564–587. https://doi.org/10.1111/j.1467-7652.2009.00497.x

High, S., Lecomte, F. J., Russell, S. J., Abell, B. M., & Oliver, J. D. (2000). Glycoprotein folding in the endoplasmic reticulum: A tale of three chaperones? *FEBS Letters*, *476*(1–2), 38–41. https://doi.org/10.1016/S0014-5793(00)01666-5

Hood, E. E., Witcher, D. R., Maddock, S., Meyer, T., Baszczynski, C., Bailey, M., … Howard, J. A. (1997). Commercial production of avidin from transgenic maize: Characterization of transformant, production, processing, extraction and purification. *Molecular Breeding*, *3*, 291–306.

Jaganathan, D., Ramasamy, K., Sellamuthu, G., Jayabalan, S., & Venkataraman, G. (2018). CRISPR for crop improvement: An update review. *Frontiers in Plant Science*, *9*, Article 985. https://doi.org/10.3389/fpls.2018.00985

Janssen, B., & Gardner, R. C. (1990). Localized transient expression of GUS in leaf discs following ocultivation with *Agrobacterium*. *Plant Molecular Biology*, *14*(1), 61–72. https://doi.org/10.1007/bf00015655

Kaiser, J. (2008). Is the drought over for pharming? *Science*, *320*(5875), 473–475. https://doi.org/10.1126/science.320.5875.473

Kamionka, M. (2011). Engineering of therapeutic proteins production in *Escherichia coli*. *Current Pharmaceutical Biotechnology*, *12*(2), 268–274. https://doi.org/10.2174/138920111794295693

Kapila, J., Rycke, R. D., Montagu, M. V., & Angenon, G. (1997). An *Agrobacterium*-mediated transient gene expression system for intact leaves. *Plant Science*, *122*(1), 101–108. https://doi.org/10.1016/S0168-9452(96)04541-4

Kay, R., Chan, A., Daly, M., & McPherson, J. (1987). Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. *Science*, *236*(4806), 1299–1302. https://doi.org/10.1126/science.236.4806.1299

Kowarik, M., Numao, S., Feldman, M. F., Schulz, B. L., Callewaert, N., Kiermaier, E., … Aebi, M. (2006). N-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase. *Science*, *314*(5802), 1148–1150. https://doi.org/10.1126/science.1134351

Kozlowski, S., & Swann, P. (2006). Current and future issues in the manufacturing and development of monoclonal antibodies. *Advanced Drug Delivery Reviews*, *58*(5–6), 707–722. https://doi.org/10.1016/j.addr.2006.05.002

Lagassé, H. A. D., Alexaki, A., Simhadri, V. L., Katagiri, N. H., Jankowski, W., Sauna, Z. E., & Kimchi-Sarfaty, C. (2017). Recent advances in (therapeutic protein) drug development, *F1000Research*, *6*, 113.

Lau, O. S., & Sun, S. S. M. (2009). Plant seeds as bioreactors for recombinant protein production. *Biotechnology Advances*, *27*(6), 1015–1022. https://doi.org/10.1016/j.biotechadv.2009.05.005

Le Mauff, F., Mercier, G., Chan, P., Burel, C., Vaudry, D., Bardor, M., … Landry, N. (2015). Biochemical composition of haemagglutinin-based influenza virus-like particle vaccine produced by transient expression in tobacco plants. *Plant Biotechnology Journal*, *13*, 717–725. https://doi.org/10.1111/pbi.12301

Lico, C., Santi, L., Twyman, R. M., Pezzotti, M., & Avesani, L. (2012). The use of plants for the production of therapeutic human peptides. *Plant Cell Reports*, *31*(3), 439–451. https://doi.org/10.1007/s00299-011-1215-7

Lim, Y., Wong, N. S. C., Lee, Y. Y., Ku, S. C. Y., Wong, D. C. F., & Yap, M. G. S. (2010). Engineering mammalian cells in bioprocessing–current achievements and future perspectives. *Biotechnology and Applied Biochemistry*, *55*(4), 175–189. https://doi.org/10.1042/BA20090363

Love, A. J., Chapman, S. N., Matic, S., Noris, E., Lomonossoff, G. P., & Taliansky, M. (2012). *In planta* production of a candidate vaccine against bovine papillomavirus type 1. *Planta*, *236*(4), 1305–1313. https://doi.org/10.1007/s00425-012-1692-0

Ma, J.-K.-C., Drake, P. M. W., & Christou, P. (2003). The production of recombinant pharmaceutical proteins in plants. *Nature Reviews Genetics*, *4*(10), 794–805. https://doi.org/10.1038/nrg1177

Manivasakam, P., Aubrecht, J., Sidhom, S., & Schiestl, R. H. (2001). Restriction enzymes increase efficiencies of illegitimate DNA integration but decrease homologous integration in mammalian cells. *Nucleic Acids Research*, *29*(23), 4826–4833. https://doi.org/10.1093/nar/29.23.4826

Margolin, E., Chapman, R., Williamson, A., Rybicki, E. P., & Meyers, A. E. (2018). Production of complex viral glycoproteins in plants as vaccine immunogens. *Plant Biotechnology Journal*, *16*, 1531–1545. https://doi.org/10.1111/pbi.12963

Marsian, J., & Lomonossoff, G. P. (2016). Molecular pharming – VLPs made in plants. *Current Opinion in Biotechnology*, *37*, 201–206. https://doi.org/10.1016/j.copbio.2015.12.007

Meyer, H. P., Brass, J., Jungo, C., Klein, J., Wneger, J., & Mommers, R. (2008). An emerging star for therapeutic and catalytic protein production. *Bioprocess International*, 10–21.

Miki, D., Zhang, W., Zeng, W., Feng, Z., & Zhu, J.-K. (2018). CRISPR/Cas9-mediated gene targeting in Arabidopsis using sequential transformation. *Nature Communications*, *9*(1), Article 1967. https://doi.org/10.1038/s41467-018-04416-0.

Montero-Morales, L., & Steinkellner, H. (2018). Advanced plant-based glycan engineering. *Frontiers in Bioengineering and Biotechnology*, *6*, 81.

Moustafa, K., Makzhoum, A., & Trémouillaux-Guiller, J. (2016). Molecular farming on rescue of pharma industry for next generations. *Critical Reviews in Biotechnology*, *36*(5), 840–850. https://doi.org/10.3109/07388551.2015.1049934

Obembe, O. O., Popoola, J. O., Leelavathi, S., & Reddy, S. V. (2011). Advances in plant molecular farming. *Biotechnology Advances*, *29*(2), 210–222. https://doi.org/10.1016/j.biotechadv.2010.11.004

Pavlou, A. K., & Reichert, J. M. (2004). Recombinant protein therapeutics - success rates, market trends and values to 2010. *Nature Biotechnology*, *22*(12), 1513–1519. https://doi.org/10.1038/nbt1204-1513

Perrin, Y., Vaquero, C., Gerrard, I., Sack, M., Drossard, J., Stöger, E., … Fischer, R. (2000). Transgenic pea seeds as bioreactors for the production of a single-chain Fv fragment (scFV) antibody used in cancer diagnosis and therapy. *Molecular Breeding*, *6*(4), 345–352.

Rader, R. (2008). Expression systems for process and product improvement. *BioProcess International*, 4–8.

Rader, R. A. (2012). FDA Biopharmaceutical Product Approvals and Trends: 2012. *BioProcess International*, *11*, 18–27.

Raju, T. S., Briggs, J. B., Borge, S. M., & Jones, A. J. (2000). 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*(5), 477–486. https://doi.org/10.1093/glycob/10.5.477

Rybicki, E. P. (2010). Plant-made vaccines for humans and animals. *Plant Biotechnology Journal*, *8*(5), 620–637. https://doi.org/10.1111/j.1467-7652.2010.00507.x

Sahdev, S., Khattar, S. K., & Saini, K. S. (2008). Production of active eukaryotic proteins through bacterial expression systems: A review of the existing biotechnology strategies. *Molecular and Cellular Biochemistry*, *307*(1–2), 249–264.

Sethuraman, N., & Stadheim, T. A. (2006). Challenges in therapeutic glycoprotein production. *Current Opinion in Biotechnology*, *17*(4), 341–346. https://doi.org/10.1016/j.copbio.2006.06.010

Shadid, N., & Daniell, H. (2016). Plant-based oral vaccines against zoonotic and non-zoonotic diseases. *Plant Biotechnology Journal*, *14*, 2079–2099. https://doi.org/10.1111/pbi.12604

Shih, S.-M.-H., & Doran, P. M. (2009). Foreign protein production using plant cell and organ cultures: Advantages and limitations. *Biotechnology Advances*, *27*(6), 1036–1042. https://doi.org/10.1016/j.biotechadv.2009.05.009

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., … Sugiura, M. (1986). The complete nucleotide sequence of the tobacco chloroplast genome: Its gene organization and expression. *The EMBO Journal*, *5*(9), 2043–2049. https://doi.org/10.1002/j.1460-2075.1986.tb04464.x

Shoji, Y., Chichester, J. A., Jones, M., Manceva, S. D., Damon, E., Mett, V., … Yusibov, V. (2011). Plant-based rapid production of recombinant subunit hemagglutinin vaccines targeting H1N1 and H5N1 influenza. *Human Vaccines*, *7*, 41–50. https://doi.org/10.4161/hv.7.0.14561

Staub, J. M., Garcia, B., Graves, J., Hajdukiewicz, P. T. J., Hunter, P., Nehra, N., … Russell, D. A. (2000). High-yield production of a human therapeutic protein in tobacco chloroplasts. *Nature Biotechnology*, *18*(3), 333–338. https://doi.org/10.1038/73796

Stöger, E., Vaquero, C., Torres, E., Sack, M., Nicholson, L., Drossard, J., … Fisher, R. (2000). Cereal crops as viable production and storage systems for pharmaceutical scFv antibodies. *Plant Molecular Biology*, *42*(4), 583–590. https://doi.org/10.1023/a:1006301519427

Strasser, R. (2016). Plant protein glycosylation. *Glycobiology*, *26*(9), 926–939. https://doi.org/10.1093/glycob/cww023

Su, J., Zhu, L., Sherman, A., Wang, X., Lin, S., Kamesh, A., … Daniell, H. (2015). Low cost industrial production of coagulation factor IX bioencapsulated in lettuce cells for oral tolerance induction in hemophilia B. *Biomaterials*, *70*, 84–93. https://doi.org/10.1016/j.biomaterials.2015.08.004

Tekoah, Y., Shulman, A., Kizhner, T., Ruderfer, I., Fux, L., Nataf, Y., … Shaaltiel, Y. (2015). Large-scale production of phamaceutical proteins in plant cell culture - the protalix experience. *Plant Biotechnology Journal*, *13*, 1199–1208. https://doi.org/10.1111/pbi.12428

Thomas, B., Deynze, A. V., & Bradford, K. (2002). *Production of therapeutic proteins in plants* (pp. 1–12). Oakland: University of California.

Tregoning, J. S. (2003). Expression of tetanus toxin fragment C in tobacco chloroplasts. *Nucleic Acids Research*, *31*(4), 1174–1179. https://doi.org/10.1093/nar/gkg221

Twyman, R. M., Stoger, E., Schillberg, S., Christou, P., & Fischer, R. (2003). Molecular farming in plants: Host systems and expression technology. *Trends in Biotechnology*, *21*(12), 570–578. https://doi.org/10.1016/j.tibtech.2003.10.002

van Dijk, M., & van de Winkel, J. (2001). Human antibodies as next generation therapeutics. *Current Opinion in Chemical Biology*, *5*(4), 368–374. https://doi.org/10.1016/S1367-5931(00)00216-7

Vaquero, C., Sack, M., Schuster, F., Finnern, R., Drossard, J., Schumann, D., … Fischer, R. (2002). A carcinoembryonic antigen-specific diabody produced in tobacco. *The FASEB Journal*, *16*(3), 408–410. https://doi.org/10.1096/fj.01-0363fje

Varki, A. (2009). Multiple changes in sialic acid biology during human evolution. *Glycoconjugate Journal*, *26*(3), 231–245. https://doi.org/10.1007/s10719-008-9183-z

Vasil, I. K. (2008). A history of plant biotechnology: From the cell theory of Schleiden and Schwann to biotech crops. *Plant Cell Reports*, *27*(9), 1423–1440. https://doi.org/10.1007/s00299-008-0571-4

Veerapen, V. P., van Zyl, A. R., Rybickia, E. P., & Meyersa, A. E. (2018). Transient expression of heat- and acid-resistant foot-and-mouth disease virus P1–2A mutants in *Nicotiana benthamiana*. *Virus Research*, *256*, 45–49. https://doi.org/10.1016/j.virusres.2018.08.004

Verma, D., Samson, N. P., Koya, V., & Daniell, H. (2008). A protocol for expression of foreign genes in chloroplasts. *Nature Protocols*, *3*(4), 739–758. https://doi.org/10.1038/nprot.2007.522

Verma, R., Boleti, E., & George, A. J. (1998). Antibody engineering: Comparison of bacterial, yeast, insect and mammalian expression systems. *Journal of Immunological Methods*, *216*(1–2), 165–181. https://doi.org/10.1016/S0022-1759(98)00077-5

Vézina, L.-P., Faye, L., Lerouge, P., D'Aoust, M.-A., Marquet-Blouin, E., Burel, C., … Gomord, V. (2009). Transient co-expression for fast and high-yield production of antibodies with human-like N-glycans in plants. *Plant Biotechnology Journal*, *7*(5), 442–455. https://doi.org/10.1111/j.1467-7652.2009.00414.x

Walsh, G. (2010). Biopharmaceutical benchmarks 2010. *Nature Biotechnology*, *28*(9), 917–924. https://doi.org/10.1038/nbt0910-917

Wang, P., Zhang, J., Sun, L., Ma, Y., Xu, J., Liang, S., … Zhang, X. (2018). High efficient multisites genome editing in allotetraploid cotton (*Gossypium hirsutum*) using CRISPR/Cas9 system. *Plant Biotechnology Journal*, *16*, 137–150. https://doi.org/10.1111/pbi.12755

Ward, B. J., Landry, N., Trepanier, S., Mercier, G., Dargis, M., Couture, M., … Vézina, L. P. (2014). Human antibody response to N-glycans present on plant-made influenza virus-like particle (VLP) vaccines. *Vaccine*, *32*, 6098–6106. https://doi.org/10.1016/j.vaccine.2014.08.079

Watson, J., Koya, V., Leppla, S. H., & Daniell, H. (2004). Expression of *Bacillus anthracis* protective antigen in transgenic chloroplasts of tobacco, a non-food/feed crop. *Vaccine*, *22*(31–32), 4374–4384. https://doi.org/10.1016/j.vaccine.2004.01.069

Wen, A. M., Shukla, S., Saxena, P., Aljabali, A. A. A., Yildiz, I., Dey, S., … Steinmetz, N. F. (2012). Interior engineering of a viral nanoparticle and its tumor homing properties. *Biomacromolecules*, *13*, 3990–4001. https://doi.org/10.1021/bm301278f

Zhang, B., Shanmugaraj, B., & Daniell, H. (2017). Expression and functional evaluation of biopharmaceuticals made in plant chloroplasts. *Current Opinion in Chemical Biology*, *31*, 17–23. https://doi.org/10.1016/j.cbpa.2017.02.007

Zhang, J., Li, Y., Jin, J., Chen, Q., Xie, X., Wang, Z., … Yang, J. (2017). Recent advances in tobacco chloroplast genetic engineering. *Tobacco Science & Technology Journal*, *50*(6), 88–98.

How to cite this article: Burnett MJB, Burnett AC. Therapeutic recombinant protein production in plants: Challenges and opportunities. *Plants, People, Planet*. 2019;00:1–12. https://doi.org/10.1002/ppp3.10073