Review article
Green giant—a tiny chloroplast genome with mighty power to produce high-value proteins: history and phylogeny

Henry Daniell1,* (Tel +1 215 746 2563; fax 1-215-898-3695; email hdaniell@upenn.edu), Shuangxia Jin2, Xin-Guang Zhu3, Matthew A. Gitzendanner4, Douglas E. Soltis5,6 and Pamela S. Soltis6

1Department of Basic and Translational Sciences, School of Dental Medicine, University of Pennsylvania, Philadelphia, PA, USA
2National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, China
3State Key Laboratory for Plant Molecular Genetics and Center of Excellence for Molecular Plant Sciences, Chinese Academy of Sciences, Shanghai, China
4Department of Biology, University of Florida, Gainesville, FL, USA
5Florida Museum of Natural History and Department of Biology, University of Florida, Gainesville, FL, USA
6Florida Museum of Natural History, University of Florida, Gainesville, FL, USA

Received 14 December 2020; revised 11 January 2021; accepted 16 January 2021.
*Correspondence (Tel +1 215 746 2563; fax 1-215-898-3695; email hdaniell@upenn.edu)

Keywords: Biopharmaceuticals, chloroplast genome structure, evolution, carbon capture, COVID-19, enzymes, plastid genetic engineering.

Summary
Free-living cyanobacteria were entrapped by eukaryotic cells ~2 billion years ago, ultimately giving rise to chloroplasts. After a century of debate, the presence of chloroplast DNA was demonstrated in the 1960s. The first chloroplast genomes were sequenced in the 1980s, followed by ~100 vegetable, fruit, cereal, beverage, oil and starch/sugar crop chloroplast genomes in the past three decades. Foreign genes were expressed in isolated chloroplasts or intact plant cells in the late 1980s and stably integrated into chloroplast genomes, with typically maternal inheritance shown in the 1990s. Since then, chloroplast genomes conferred the highest reported levels of tolerance or resistance to biotic or abiotic stress. Although launching products with agronomic traits in important crops using this concept has been elusive, commercial products developed include enzymes used in everyday life from processing fruit juice, to enhancing water absorption of cotton fibre or removal of stains as laundry detergents and in dye removal in the textile industry. Plastid genome sequences have revealed the framework of green plant phylogeny as well as the intricate history of plastid genome transfer events to other eukaryotes. Discordant historical signals among plastid genes suggest possible variable constraints across the plastome and further understanding and mitigation of these constraints may yield new opportunities for bioengineering. In this review, we trace the evolutionary history of chloroplasts, status of autonomy and recent advances in products developed for everyday use or those advanced to the clinic, including treatment of COVID-19 patients and SARS-CoV-2 vaccine.

Introduction
Chloroplasts have sustained life on this planet for one billion years by providing carbohydrates, amino acids, lipids for human nutrition and O2 to breathe through the process of photosynthesis. Chloroplasts synthesize the most abundant protein on Earth (RuBisCO) and utilize CO2, thereby playing a key role in reducing atmospheric CO2 and moderating global warming. Although non-Mendelian inheritance was recognized in early 1900s (Baur, 1909; Correns, 1909), direct evidence for chloroplast DNA was not established until 1960s (Kirk, 1963; Ris and Plaut, 1962). In the past three decades, chloroplast genomes have been genetically modified to introduce agronomic traits mostly in model systems or to produce therapeutic products or enzymes used in everyday life. The typical maternal inheritance (with rare paternal or biparental inheritance) of chloroplast genomes offers gene containment when foreign genes are integrated into chloroplast genomes. The large copy number of chloroplast genomes (up to 10,000 in each plant cell) offers high levels of expression of introduced foreign genes. Site-specific integration of foreign genes reduces the number of transgenic lines to be evaluated. In addition, the lack of gene silencing or position effects in chloroplast genomes offers uniform expression of foreign genes. Thus, modern-day chloroplasts could play additional major roles in human health and well-being well beyond providing food, feed, fuel and oxygen.

Integration of foreign genes (>300) into chloroplast genomes resulted in their highest levels of expression and conferred the highest level of tolerance or resistance reported to date. Insulin was expressed in chloroplasts up to 70% of leaf protein or higher than RubisCO (Ruhlman et al., 2010), and a high-value polymer, pHBA, was expressed up to 26.5% of leaf dry weight (Vitican et al., 2004). The Cry2A operon produced folded Bt crystals within chloroplasts (DeCosa et al., 2001). Chloroplast expression...
of Tic40 triggered the nucleus to make all inner membrane proteins, resulting in 8-20 layers of the inner envelope membrane (Singh et al., 2008). Although these are early examples of using very small human proteins or genes of bacterial origin, most of the challenging problems with engineered plastids remained unaddressed until recently. Failure to adequately express human or viral genes in chloroplasts has been addressed recently by developing new algorithms that relied on a large number of sequenced chloroplast genomes (Kwon et al., 2016; Kwon et al., 2018). When the largest human blood protein (FVIII, 185 kDa monomer, with pentamer assembly) was expressed in the chloroplast, the level reached industry expectations for commercial scale production and clinical evaluation (Kwon et al., 2018). Codon optimised expression of angiotensin-converting enzyme 2 not only helped to treat pulmonary hypertension, but has also now been tested in the clinic to treat COVID-19 patients (Daniell, 2020), as discussed in depth below. Codon-optimized expression of a novel human insulin-like growth factor in lettuce chloroplasts promoted musculoskeletal cell proliferation, differentiation and diabetic bone fracture healing (Park et al., 2020). The expression level of enzymes in chloroplasts and the efficacy exceeded current products made in fermentation systems using almost all evaluation criteria, without any need for purification (Daniell et al., 2019c; Kumari et al., 2019). Most importantly, a selectable antibiotic resistance gene has been removed from transplastomic chloroplast genomes in edible crops (Daniell et al., 2019b; Kumari et al., 2019; Park et al., 2020). Production of protein drugs in cGMP facilities, long-term stability of drugs in freeze-dried plant cells (Daniell, 2020; Daniell et al., 2019b; Daniell et al., 2019c; Park et al., 2020; Su et al., 2015), evaluation of drug substance by toxicology and pharmacokinetic studies (Daniell et al., 2019c) and recent FDA approval of PDs orally delivered peanut cells to treat allergy (Tilles and Petroni, 2018; Vickery et al., 2018) all represent major recent advances in this field. Based on these recent advances, the next decade should be particularly promising for the common use of proteins made in chloroplasts to enhance human health and well-being.

This review offers a brief overview of the history, changes in chloroplast autonomy, the use of plastid DNA sequence data to reconstruct plant evolutionary history, as well as recent advances and future use of chloroplast genomes for biotechnology applications. This summary is not intended to provide a complete list of all chloroplast genomes sequenced or foreign genes expressed but our goal is to focus on a few selected products advanced beyond laboratories for daily use or in clinical development. In addition, philosophical questions on the loss of chloroplast genes during evolution and status of chloroplast autonomy are discussed.

Female dominance and transgene containment

In 1909, Baur and Correns described the revolutionary concept of non-Mendelian inheritance factors located outside the nucleus in the protoplasm (Baur, 1909; Correns, 1909; Table 1). In crosses between variegated, yellow and green Mirabilis jalapa plants, branches always gave rise to the same colour seedlings, without any Mendelian segregation (Correns, 1909). Studies in the following decades showed different mechanisms for exclusion of male chloroplast DNA during sexual fusion. The generative cells divide unequally during pollen formation and do not receive any plastids. In angiosperms, each generative cell forms two sperm cells, one of which fuses with a female gamete to form a zygote and the other of which unites with polar nuclei to produce endosperm, that is the process referred to as ‘double fertilization’. In this process, male chloroplasts or DNA is specifically excluded or degraded, assuring maternal inheritance of the chloroplast genome in both the vast majority of angiosperms and other land plants as well, with a few noteworthy exceptions (Daniell, 2002; Daniell, 2007; Hagemann and Schroeder, 1989).

Role of maternal inheritance of chloroplast genomes gained greater attention when foreign genes were introduced into chloroplast genomes. Transgene escape via pollen and the possibility of weedy relatives capturing this valuable trait have been a major concern in nuclear transgenic crops. Therefore, integration of the herbicide resistance gene via the chloroplast genome and maternal inheritance of transgenes was considered a major accomplishment, and this invention was featured on the cover of Nature Biotechnology (Daniell et al., 1998). Subsequently, maternal inheritance of numerous foreign genes integrated into the chloroplast genome has been documented (Daniell et al., 2016a; Daniell et al., 2016b; Jin and Daniell, 2015; Table 1).

In addition to transgene containment via maternal inheritance, chloroplast engineering produces products in leaves, facilitating their harvest before appearance of any reproductive structures, thus offering complete containment. Therefore, field tests of products were conducted a decade ago for biopharmaceuticals expressed in chloroplasts (Arlen et al., 2007) (Table 1). More recently, USDA-APHIS certifies that ‘transplastomic lines do not fit the definition of a regulated article under USDA-APHIS regulations 7 CFR part 340, because there are no plant pest components’ (Kwon and Daniell, 2015). These advantages should facilitate field production of products engineered via the chloroplast genome.

Entrapment and semi-autonomous state of chloroplasts

The concept of cyanobacteria as evolutionary precursors of chloroplasts was proposed a century ago (Mereschkowsky, 1905), and a cyanobacterial origin of all plastids is now well established (Delwiche et al., 1995). About 1.5–2 billion years ago, free-living cyanobacteria were entrapped by early eukaryotic cells. This entrapment process has been experimentally demonstrated in laboratories through the uptake of algal chloroplasts by plant protoplasts or of green chloroplasts by albino protoplasts, and regeneration of green or variegated plants (Bonnett, 1976; Bonnett and Eriksson, 1974). Subsequent to this ancient entrapment event, there was a massive transfer of cyanobacterial genes to the nuclear genome over the long course of evolutionary history of green plants. In Arabidopsis, 18% of the nuclear genes originated from the ancestral plastid genome (Martin et al., 2002). Therefore, most of the plastid proteins are nuclear-encoded and genetically dependent on the host cell, and chloroplasts are no longer obligate endosymbionts but organelles (Cavalier-Smith, 1985).

There has been an enormous reduction in gene content from the ancestral cyanobacteria to the plastid genomes found in photosynthetic eukaryotes. The genome of the cyanobacterium Synechocystis PCC 6803 is 3,573 kbp and contains ~3,200 genes, and the plastid genome in the red alga Porphyra purpurea is only 191 kbp, with ~250 genes (Reith and Munholland, 1995). Plastids of red algae and glaucocystophytes encode genes for several biosynthetic pathways, nitrogen assimilation and metabolic regulation, in addition to genes involved in protein synthesis and photosynthesis (Delwiche, 1999). Indeed, the plastid genome of the glaucophyte Cyanophora paradoxa contains genes for
### Table 1: Milestones in plant chloroplast genome and genetic engineering/biotechnology (first reports and subsequent validation by further research). First reports on chloroplast transformation in different crop species, without useful products (only aadA gene), although important milestones, are not listed here due to inadequate follow-up studies and space limitations

| Year(s) | Milestone Developments | Author/group/References |
|---------|------------------------|-------------------------|
| 1909    | Non-Mendelian inheritance | Baur (1909), Correns (1909) |
| 1929    | 'Plastome' as hereditary factor | Renner (1929) |
| 1962    | Chlamydomonas Cp DNA—visual evidence | Kiris and Plaut (1962) |
| 1963    | Broadbean cp DNA—visual evidence | Kirk (1963), Leff et al (1963) |
| 1974-76 | Uptake of isolated chloroplasts by protoplasts and plant regeneration | Bonnett and Eriksson (1974), Bonnett (1976) |
| 1983-86 | Synthesis of thylakoids, macro-grana functional PSI, PSII in isolated chloroplasts | Daniell et al. (1983, 1984, 1986) |
| 1986    | Complete sequence of the first plant chloroplast genome (tobacco) | Shinozaki et al. (1986) |
| 1987    | First foreign genes (chloramphenicol acetyltransferase, B-lactamase) expressed in isolated chloroplasts— | Daniell and McFadden (1987) |
| 1990    | First chloroplast vectors using the psbA regulatory sequences subsequently used in large majority of transgene expression studies in chloroplasts | Daniell et al. (1990, 1991) |
| 1990    | First foreign gene expression in plant chloroplasts (chloramphenicol acetyl transferase, B-glucuronidase) using the gun powder or helium gene gun—the most reproducible gene delivery system for chloroplast transformation | Daniell et al. (1990, 1991), Ye et al. (1990) |
| 1991    | The aadA gene as chloroplast selectable marker—subsequently used in most transgenes integrated into chloroplast genomes | Goldschmidt-Clermont (1991) |
| 1993    | The aadA gene integration into the tobacco chloroplast genome in the large single-copy region | Svab and Maliga (1993) |
| 1995    | Cry1A gene expression in chloroplasts to confer resistance to insects, followed by subsequent reports in other genes | McBride et al. (1995), Kota et al. (1999), Chakrabarti et al. (2006) |
| 1998    | First transgene integration into the inverted Repeat regions of the chloroplast genome, subsequently confirmed by other labs to significantly enhance transgene expression by doubling the copy number, accelerate integration due to copy correction and used in large majority of transgene expression studies | Daniell et al. (1998), Krichevsky et al. (2010) |
| 1998    | Maternal inheritance of herbicide resistance gene, reported in most subsequent studies for this or several other traits | Daniell et al. (1998), Lutz et al. (2001), Ye et al. (2001) |
| 2000    | Marker-free transplastomic tobacco from the large single-copy region, followed by excision in soybean transplastome | Lamtham and Day (2000), Dufournantel et al. (2007) |
| 2000    | First biopolymer, biopharmaceutical subsequently, numerous biopharmaceuticals have been expressed in chloroplasts | Guda et al. (1999); Staub et al. (2000) |
| 2001    | First foreign operon (Cry2A) and Bt crystals in chloroplasts, several operons were expressed subsequently | DeCosa et al. (2001), Malhotra et al. (2016), Fuentes et al. (2016) |
| 2001    | First vaccine antigen (CTB) in chloroplasts followed by several vaccine antigens challenged with toxins or pathogens in animal models against anthrax, cholera, dengue, HIV, malaria, plague, polio, tetanus, etc. | Daniell et al. (2001), Koya et al. (2005), Arlen et al. (2008), Davoodi-Semiromi et al. (2010), Gonzalez-Rabade et al. (2011), Chan et al. (2016), van Erde et al. (2019), Daniell et al. (2019b), Tregoning et al. (2005) |
| 2001    | First antimicrobial peptide followed by several other studies | DeGray et al. (2001), Oey et al. (2009), Gupta et al. (2015), Lee et al. (2011) |
| 2003    | First cell wall degrading enzyme in chloroplasts followed by enzyme cocktails for plant biomass hydrolysis, detergent or textile applications | Leelavathee et al. (2003), Verma et al. (2010), Agrawal et al. (2011), Petersen and Bock (2011), Daniell et al. (2019), Kumari et al. (2019) |
| 2003    | First phyto remediation using chloroplasts followed by several subsequent studies | Ruiz et al. (2003, 2011), Hussein et al. (2007) |
| 2004    | First engineering via somatic embryogenesis and non-green carrot plastids, conferred salt-tolerance, followed by demonstration of soybean with agronomic traits | Kumar et al. (2004), Dufournantel et al. (2004), Dufournantel et al. (2007) |
| 2004    | First metabolic engineering in chloroplasts followed by several other products | Vitanen et al. (2004), Harada et al. (2014), Apel and Bock (2009), Pasrecek et al. (2016) |
| 2004    | First RuBisCO nuclear SSU expression in chloroplasts followed by carboxysomes | Dhingra et al. (2004), Shanwood et al. (2008), Long et al. (2018) |
| 2005    | Complete chloroplast genome sequences of common crops—Potato, tomato, soybean, carrot, coffee, grape, orange, cotton, cassava, cocoa, lettuce, etc. | Saski et al. (2005), Daniell et al. (2006, 2008), Jansen et al. (2006, 2008), Rauher et al. (2006), Samson et al. (2007), Lee et al. (2006), Ruhlman et al. (2006) |
biosynthesis of the peptidoglycan wall, and this species retains a cyanobacterial cell wall (Löffelhardt and Bohnert, 1994). Loss of the peptidoglycan wall during the evolution of land plants is a significant irreversible step towards reducing plastid survival outside the plant cell. Reduction of plastid genome size continues (up to 113 kbp) with loss of both the inverted repeat region and ndh genes in land plants (Sanderson et al., 2015). Most land plant plastid genomes contain 110–130 genes, with ~80 genes coding for proteins involved in photosynthesis and other processes. In parasitic plants, photosynthetic genes are lost, but the chloroplast genomes retain genes essential for protein synthesis and origin of DNA replication (Banerjee and Stefanović, 2019). Despite the heavy loss of genes from the ancestral cyanobacterial genome, isolated chloroplasts still retain the ability to perform protein synthesis because they contain >50% of the total ribosomal complement of photosynthetic cells, DNA, DNA polymerase, RNA polymerase and tRNAs (Ellis, 1977). Use of S19 methionine and specific protein synthesis inhibitors for chloroplast (chloramphenicol) and cytosolic (cloheximide) ribosomes made it possible to distinguish proteins synthesized in each compartment (Ellis, 1977). Isolated intact chloroplasts were shown to synthesize proteins, capable of photosystem I activity with cyclic phosphorylation (Daniell et al., 1983) and form macrograna (Rebeiz et al., 1984) with photosystem II activities (Daniell et al., 1984; Daniell and Sarojini, 2019) during greening in vitro. It is remarkable that protein synthetic capacity is retained in parasitic plants when all photosynthetic genes were lost.

### Plastome—a century of progress

Renner coined the term ‘plastome’ to describe plastid hereditary factors (Renner, 1929, 1934) (Table 1). Although the presence of DNA within plastids was debated for decades, Ris and Plaut (1962) were the first to convincingly demonstrate the presence of DNA in *Chlamydomonas* chloroplasts using electron micrographs with 25 Å microfibril size, sensitivity to nuclease digestion and Feulgen reaction and yellow-green fluorescence of acridine orange staining. Subsequently, DNA was shown in broad bean (Kirk, 1963) and other chloroplasts and was referred to as ‘satellite DNA’ (Leff et al., 1963) (Table 1).

Two decades after discovery of organellar DNA, the complete chloroplast genome of the first crop plastome genome (tobacco) was published by Shinozaki et al in 1986 (Table 1). Only a few crop chloroplast genome sequences were published in the following decade, and this paucity of data was a major limitation in engineering crop plastid genomes. The misconception was that chloroplast genome sequences are highly conserved among crop plastid genomes. However, ~50% of the plastid genome contains coding sequences which are highly conserved, but the intergenic sequences that are essential for transgene integration or that contain regulatory sequences are not conserved. Among species of Solanaceae, only four of >150 intergenic sequences are conserved, including the trnL/trnA spacer region used in our chloroplast vectors (Daniell et al., 2006). Not even a single spacer region is conserved among sequenced grass

| Year (s) | Milestone Developments | Author/group/References |
|---------|------------------------|-------------------------|
| 2006    | First biopharmaceutical in edible crop – lettuce followed by several protein drugs | Ruhlman et al. (2007, 2010), Boyhan and Daniell (2011), Su et al (2015), Kwon et al (2016, 2018), Park et al (2020), Daniell et al (2020) |
| 2007    | First field production of cp biopharmaceuticals followed by enzyme field production | Arlen et al. (2007), Schmidt et al. (2019) |
| 2008    | First membrane protein expressed in chloroplasts increased inner envelope up to 19 layers, followed by 8 layers with TMT expression | Singh et al. (2008), Jin and Daniell (2014) |
| 2010    | First oral tolerance induction to treat allergies to protein drugs, followed by FVIII, Pompe | Ruhlman et al (2007), Verma et al. (2010), Sherman et al (2014), Kwon et al. (2018), Su et al. (2015) |
| 2015    | First DS RNA expression in chloroplasts confers protection against insects, followed by subsequent studies | Zhang et al. (2015), Jin et al. (2015), He et al. (2020) |
| 2015    | CGMP production and evaluation of protein drugs made in chloroplasts | Su et al. (2015), Park et al (2020), Daniell et al (2020) |
| 2016    | Evaluation of chloroplast polio vaccine by CDC, FDA, Gates Foundation team | Chan et al. (2016), Xiao and Daniell (2017), Daniell et al. (2019a) |
| 2016    | Codon optimization using >130 chloroplast genomes and expression of the largest human protein in lettuce chloroplasts | Kwon et al. (2016, 2018), Chan et al (2016), Park et al (2019), Daniell et al (2020) |
| 2017    | Lettuce chloroplast FXI protects haemophilia dogs from anaphylaxis -Novo Nordisk industry evaluation | Herzog et al. (2017) |
| 2019    | First marker-free enzymes/biopharmaceutical in lettuce chloroplasts, integrated into the inverted repeat region, followed by other reports | Daniell et al. (2019c), Kumari et al (2019), Park et al (2020), Daniell et al (2020) |
| 2019    | Commercial launch of chloroplast enzyme products for textile/detergent—by PhyloZyme | Daniell et al. (2019c), Kumari et al (2019) |
| 2019    | First investigational New Drug studies of biopharmaceutical made in chloroplasts and advanced to human clinical trials –COVID-19 | Daniell et al (2020), Daniell et al (2020) |
| 2019    | First chloroplast transformation using carbon nanotubes | Kwak et al. (2019) |
| 2020    | First Cas9/gRNA chloroplast genome editing, followed by subsequent studies; | Yoo et al. (2020), Zhan et al. (2019) |
chloroplast genomes (Sasaki et al., 2007). Transgene cassettes are introduced into the chloroplast genome intergenic spacer regions, using native genes to facilitate homologous recombination. However, when tobacco chloroplast genome flanking sequences were inserted into the lettuce chloroplast genome, unique nucleotides were completely eliminated or modified to achieve 100% homologous recombination, dramatically reducing transformation efficiency (Ruhmlan et al., 2010). Likewise, when tobacco psaA regulatory sequences (promoter, 5′; 3′ UTR) were used in lettuce chloroplasts, there was 80-97% reduction in translation when compared to the endogenous regulatory sequences (Ruhmlan et al., 2010). As a result, species-specific chloroplast vectors with endogenous genes and regulatory sequences are required for efficient foreign gene expression. Based on the work noted above, in the last decade, significant efforts were made to sequence crop chloroplast genomes that are used in everyday life including soybean (Sasaki et al., 2005), other legumes (Jansen et al., 2008), potato (Daniell et al., 2006), tomato (Daniell et al., 2006), grape (Jansen et al., 2006), coffee (Samson et al., 2007), cotton (Lee et al., 2006), orange (Bausher et al., 2006), cassava (Daniell et al., 2008), carrot (Ruhmlan et al., 2006) and cereals (Sasaki et al., 2007). For a current list of edible crop chloroplast genomes, readers are referred to FAO: http://www.fao.org/fileadmin/templates/ess/documents/world_census_of_agriculture/appendix3_r7.pdf. - and Table 2. As discussed below, chloroplast genome sequences facilitate codon optimization and offer the best regulatory sequences to enhance translation and transgene integration. Understanding the origins of economically important cultivated species facilitates breeding and prevents cross-contamination of plants used in herbal medicine. Moreover, an understanding of the diversity of chloroplast genomes, in terms of both structure and sequence, is important for developing efficient systems for genetic engineering. However, among ~3,000 cultivated crops, completely sequenced chloroplast genomes are available for fewer than 70 genera. Among these, <80 complete chloroplast genome are available in the NCBI database. However, the One Thousand Plants Transcriptome Project (1KP; onekp.com), as well as other recent efforts, have contributed over 1,000 complete or nearly complete plastid genomes to global databases, most of these from plants that are not of economic importance (Gitzendanner et al., 2018; Leebens-Mack, 2019; Li et al., 2019), hence, our understanding of plastid genomes across the Tree of Life has improved dramatically in the past decade.

**Fundamental tool in phylogenetics and evolution**

For a number of reasons (abundance, single-copy genes, lack of recombination and appropriate rate of nucleotide evolution), the plastid genome has long been the primary workshop for studies of plant phylogeny and evolution. The size and structure of the plastid genome have been remarkably conserved across land plant evolution (although intergenic spacer regions and regulatory sequences are not well conserved), in stark contrast to the enormous variation in size and structure of the plant mitochondrial genome, and this conservation has facilitated the use of both sequence data and plastome rearrangements in phylogenetic analyses. As noted above, transfer of genes from the plastome to the nuclear genome has reduced the size of the plastid genome over the course of green plant evolution, with chlorophylls having larger plastid genomes and more genes than streptomycetes, particularly land plants. There is also evidence of some plastid gene movement to the mitochondrial genome.

Phylogenetic analyses using plastid genes have been conducted across a range of divergences from the species level and to very deep levels. Particularly at deeper levels (e.g. at divergences traditionally recognized at the family level and deeper), plastid data have been of enormous value. Initial studies employed only rbcL (encoding the large subunit of RuBisCO); in a landmark study showing the utility of plastid gene sequences, a collaboration of 43 investigators provided the first DNA phylogenetic framework for seed plants based on an analysis of 499 species (Chase et al., 1993). Most recently, next-generation sequencing has enabled the sequencing of the complete plastid genome and the assembly of large phylogenetic trees across all green plants (Gitzendanner et al., 2018; Ruhfel et al., 2014); other studies of plastid loci have focused on major subclades of green plants (Li et al., 2019). Plastid phylogenetics ushered in the most the fundamental changes in our understanding of plant relationships in the past 150 years, revealing the major clades of green plants, the sister group to land plants, relationships across land plants, with a major reshaping of our understanding of moss, liverwort, fern, gymnosperms and angiosperm phylogeny. Not only have these studies resulted in a clearer understanding of evolutionary relationships, they have also prompted major new classifications for the angiosperms (APG IV 2016) and ferns (Pteridophyte.Phylogeny.Group, 2016); these are groundbreaking classifications that represent dramatic changes from anything previously published based on morphology.

In some cases, the plastid genome has exhibited enough variation to be of utility in studies at the population level and also in phylleogeographic analyses (Brunsfield et al., 2001; Soltis et al., 1997), although not on the scale observed with mtDNA in animals. Although hybridization has long been known to be a major force in plant evolution, molecular studies using plastid genes have revealed many unsuspected past hybridization events showing that hybridization is even more prevalent in plants than thought, with hundreds of documented cases of introgression of plastid genomes. Most of our current framework of green plant phylogenetic relationships is based on plastid genome sequence data, and current classifications are largely based on plastid gene phylogenetics. Only in the past few years as nuclear gene sequencing has become more routine have comparable nuclear gene topologies been generated. Importantly, there are discordances between plastid and nuclear trees, not only at shallow levels where introgression has long been detected, but also at deep levels (Stull et al., 2020; Sun, 2015), indicating putative ancient reticulation.

Studies of plastid genes and genomes have also revealed the complex history of the plastid green plant clade, with secondary and tertiary endosymbiotic events (representing the capture of photosynthetic green or red algae) occurred in other lineages, including brown algae, red algae and Euglena (Keeling, 2004; Keeling, 2010; Palmer et al., 2004). Together, this increasingly large set of plastid genes and genomes from across green plant phylogeny and other clades of photosynthetic eukaryotes provides the sequence information and resources, not only for tracing plant evolution, but also for chloroplast genetic engineering.

The technical innovations (Moore et al., 2006; Stull et al., 2013; Uribe-Convers et al., 2014) that enabled use of the entire plastome, or at least most of the ~80 protein-coding genes, as well as the four tRNA genes, typical of an angiosperm plastome, in phylogenetic analyses (Gitzendanner et al., 2018; Jansen et al., 2007; Li et al., 2019; Moore et al., 2007, 2010; Ruhfel et al.,...
| Common name       | Species                        | Genome size | Accession No. | References                                                                 |
|-------------------|--------------------------------|-------------|---------------|-----------------------------------------------------------------------------|
| Vegetables        |                                |             |               |                                                                             |
| Onion             | Allium cepa                    | 153538      | NC_024813     | von Kohn et al. (2013)                                                      |
| Sweet Pepper      | Capsicum annuum                | 156781      | NC_018552     | Jo et al. (2011)                                                            |
| Chickpea          | Cicer arietinum                | 125319      | NC_011163     | Jansen et al. (2008)                                                        |
| Broccoli          | Brassica oleracea              | 153366      | KR_233156     | Seol et al. (2017)                                                          |
| Cucumber          | Cucumis sativus                | 155293      | NC_007144     | Plader et al. (2007)                                                        |
| Carrot            | Daucus carota                  | 155911      | NC_008325     | Ruhlman et al. (2006)                                                       |
| Celery            | Apium graveolens               | 152050      | MK036045      | Zhu et al. (2019)                                                           |
| Lettuce           | Lactuca sativa                 | 152765      | NC_007578     | Jansen and Palmer (1987)                                                    |
| Pea               | Pisum sativum                  | 122169      | NC_014057     | Magee et al. (2010)                                                         |
| Kidney bean       | Phaseolus vulgaris             | 150285      | NC_009259     | Guo et al. (2007)                                                           |
| Radish            | Raphanus sativus               | 153368      | NC_024469     | Jeong et al. (2014)                                                         |
| Tomato            | Solanum lycopersicum           | 155461      | NC_007898     | Daniell et al. (2006)                                                       |
| Spinach           | Spinacia oleracea              | 150725      | NC_002202     | Schmitz-Linneweber et al. (2001)                                            |
| Mung bean         | Vigna radiata                  | 151271      | NC_013843     | Tangphasomruang et al. (2010)                                              |
| Eggplant          | Solanum melongena              | 154289      | KG_62719      | Scienctechnology et al. (2016)                                             |
| Cabbages/Turnips  | Brassica rapa                  | 153482      | DQ231548      | Li et al. (2017)                                                            |
| Globe Artichoke   | Cynara scolymus                | 152529      | KM035764      | Curci et al. (2015)                                                         |
| Asparagus         | Asparagus officinalis.          | 156699      | LN396355      | Sheng et al. (2017)                                                         |
| Chicory           | Cichorium intybus              | 15975      | MK569377      | Yang et al. (2019b)                                                         |
| Pumpkin           | Cucurbita pepo                 | 157343      | MH031787      | Zhang et al. (2018)                                                         |
| Garlic            | Allium sativum                 | 153189      | MK335928      | Huo et al. (2019)                                                           |
| Welsh onion       | Allium fistulosum              | 153162      | MK335927      | Huo et al. (2019)                                                           |
| Chive             | Allium tuberosum               | 154056      | MK335929      | Huo et al. (2019)                                                           |
| Spice Crops       |                                |             |               |                                                                             |
| Chili pepper      | Capsicum chinense              | 156807      | NC_030543.1   | Park et al. (2016)                                                          |
| Aniseed/Badian    | Illicium verum                 | 142747      | NC_034689     | Park et al. (2019)                                                          |
| Fennel            | Foeniculum vulgare             | 153628      | NC_029469     | None                                                                        |
| Nutmeg            | Torreya grandis                | 136949      | KY369757      | Mu et al. (2018)                                                            |
| Cinnamon          | Cinnamomum aromaticum          | 152754      | MN173819      | Xie et al. (2019)                                                           |
| Ginger            | Zingiber officinale            | 162621      | MH161428      | Cui et al. (2019)                                                           |
| Black pepper      | Piper nigrum                   | 161523      | NC_034692     | None                                                                        |
| Japanese pepper   | Zanthoxyllum piperitum.         | 158154      | NC_027939     | Lee et al. (2015)                                                           |
| Fruits/Melons     |                                |             |               |                                                                             |
| Pineapple         | Ananas comosus                 | 159636      | NC_026220     | Nashima (2015)                                                              |
| Papaya            | Carica papaya                  | 160100      | NC_010323     | Ming et al. (2008)                                                          |
| Wild strawberry   | Fragaria vesca                 | 155691      | NC_015206     | Shulaev et al. (2011)                                                       |
| Banana            | Musa textilis                  | 161347      | NC_022926     | Martin et al. (2013)                                                        |
| Vanilla           | Vanilla planifolia             | 148011      | NC_026778     | Lin et al. (2015)                                                           |
| Kiwifruit         | Actinidia chinensis            | 156346      | NC_026690     | Yao et al. (2015)                                                           |
| Chestnut          | Castanea mollissima            | 160799      | NC_014674     | Jansen et al. (2011)                                                        |
| Coco plum         | Chrysobalanus icaco            | 162775      | NC_024061     | Malé et al. (2014)                                                          |
| Sweet Orange      | Citrus sinensis                | 160129      | NC_008334     | Bausher et al. (2006)                                                       |
| Karaka nut        | Cynocarpus laevigata            | 159202      | NC_014807     | Atherton et al. (2010)                                                      |
| Date palm         | Phoenix dactylifera            | 158462      | NC_013991     | Yang et al. (2010)                                                          |
| Peach             | Prunus persica                 | 157790      | NC_014697     | Jansen et al. (2011)                                                        |
| Cranberry         | Vaccinium macrocarpon          | 176045      | NC_019616     | Fajardo et al. (2013)                                                       |
| Wine grape        | Vitis vinifera                 | 160928      | NC_007957     | Jansen et al. (2006)                                                        |
| Chinese pear      | Pyrus pyrifolia                | 150922      | NC_015996     | Terakami et al. (2012)                                                      |
| Cultivated Apple  | Malus domestica Borkh           | 160062      | MHS595653     | Yan et al. (2019)                                                           |
| Wild Apricot      | Prunus sibirica                | 158248      | MN078049      | Dong et al. (2020)                                                          |
| Dwarf cherry      | Cerasus humilis                | 156084      | MF405921      | Mu et al. (2018)                                                            |
| Mango             | Mangifera indica               | 157837      | NC_035239     | Zhao et al. (2019)                                                          |
| Chinese Cherry    | Prunus pseudocerasus           | 157834      | KX_255667     | Cao et al. (2018)                                                           |
| Sweet Cherry      | Cerasus avium                  | 157987      | MH_756631     | Chen et al. (2018)                                                          |
| Litchi            | Litchi chinensis               | 162524      | KY_635881     | Rabah et al. (2017)                                                         |
| Longan            | Dimocarpus longan              | 160833      | MG_214255     | Wang et al. (2017)                                                          |
| Common name   | Species                       | Genome size | Accession No. | References                                   |
|--------------|-------------------------------|-------------|---------------|----------------------------------------------|
| Avocado      | Persea americana Mill        | 152723      | KX437771      | Song et al. (2016)                           |
| Date         | Ziziphus jujuba               | 161466      | NC_030299.1   | Ma et al. (2017)                             |
| Fig          | Ficus carica L.               | 160602      | NC_035237     | Rabah et al. (2017)                          |
| Key lime     | Citrus aurantifolia          | 158983      | KJ865401.1    | Su et al. (2014)                             |
| Watermelon   | Citrullus lanatus             | 156699 - 156907 | KY_430683-KY_430693 | Shi et al. (2017)                           |
| Sweet Melon  | Cucumis melo                  | 156017      | JF412791      | Rodríguez-Moreno et al. (2011)               |
| Cereals      |                               |             |               |                                              |
| Pearl millet | Cenchrus americanus           | 140718      | NC_024171     | Mariac et al. (2014)                         |
| White fonio  | Digitaria exilis              | 140908      | NC_024176     | Mariac et al. (2014)                         |
| Barnyard grass | Echinochloa oryzicola     | 139891      | NC_024643     | Ye et al. (2014)                             |
| Buckwheat    | Fagopyrum esculentum          | 159599      | NC_010776     | Logacheva et al. (2008)                      |
| Barley       | Hordeum vulgare               | 136462      | NC_008590     | Sasaki et al. (2007)                         |
| Rice         | Oryza sativa                  | 134525      | X15901        | Hiratsuka et al. (1989)                      |
| Bread wheat  | Triticum aestivum             | 134545      | NC_002762     | Ogihara et al. (2002)                        |
| Rye          | Secale cereale                | 114843      | NC_021761     | Middleton et al. (2014)                      |
| Maize        | Zea mays                      | 140384      | NC_001666     | Maier et al. (1995)                          |
| Sorghum      | Sorghum bicolor              | 140754      | NC_008602     | Sasaki et al. (2007)                         |
| Oat           | Avena sativa                  | 135890      | NC_027468     | Saarela et al. (2015)                        |
| Adlay        | Coix lacryma-jabbi           | 140745      | FJ261955      | Kang et al. (2018)                           |
| Foxtail millet | Setaria italica            | 135516      | NC_022850     | Wang and Gao (2015)                          |
| Oil Crops    |                               |             |               |                                              |
| Canola       | Brassica napus                | 152860      | NC_016734     | Hu et al. (2010)                             |
| Soybean      | Glycine max                   | 152218      | NC_007942     | Sasaki et al. (2005)                         |
| Sunflower    | Helianthus annuus             | 151104      | NC_007977     | Timme et al. (2007)                          |
| Castor bean  | Ricinus communis              | 163161      | NC_016736     | Rivarola et al. (2011)                       |
| Sesame       | Sesamum indicum               | 153324      | NC_016433     | Yi and Kim (2012)                            |
| Tea oil plant | Camellia oleifera             | 156971      | NC_023084     | Shi et al. (2013)                            |
| Coconut      | Cocos nucifera                | 154731      | NC_022417     | Huang et al. (2013)                          |
| Oil palm     | Elaeis guineensis             | 156973      | NC_017602     | Uthaipasanwong et al. (2012)                 |
| Olive        | Olea europaea                 | 155862      | NC_015604     | Besnard et al. (2011)                        |
| Peanuts      | Arachis hypogaea              | 156395      | KX257487      | Yin et al. (2017)                            |
| Flax         | Linum usitatissimum           | 156721      | KY849971      | de Santana Lopes et al. (2018)               |
| Mustard      | Brassica juncea               | 153483      | NC_028272.1   | Prabhudas et al. (2015)                      |
| Niger seed   | Guizotia abyssinica           | 151762      | NC_010601.1   | Dempewolf et al. (2010)                      |
| Safflower    | Carthamus tinctorius          | 153675      | NC_030783.1   | Lu et al. (2015)                             |
| Beverage Crops |                               |             |               |                                              |
| Coffee       | Coffea arabica                | 155189      | NC_008535     | Samson et al. (2007)                         |
| Cacao tree   | Theobroma cacao               | 160604      | HK_364040     | Jansen et al. (2011)                         |
| Tea tree     | Camellia sinensis             | 157025      | MH_042531     | Meng et al. (2018)                           |
| Starch/Sugar Crops |                       |             |               |                                              |
| Potato       | Solanum tuberosum            | 155312      | DQ231562      | Chung et al. (2006)                          |
| Cassava      | Manihot esculenta             | 161453      | EU117376      | Danielli et al. (2008)                       |
| Yam          | Dioscorea polystachya         | 153243to153292 | MG267375 to MG267378 | Cao et al. (2018)                           |
| Sweet potato | Ipomoea batatas              | 161303      | NC_026703     | Yan et al. (2015)                            |
| Sugar beet   | Beta vulgaris                 | 149637      | EF534108.1    | Li et al. (2014)                             |
| Sugar cane   | Saccharum officinarum         | 141187      | MN204507      | Xu et al. (2019)                             |
| Sweet sorghum | Sorghum bicolor              | 141266      | NC53562       | Yang et al. (2019a)                          |
| Medicinal plants |                               |             |               |                                              |
| Guan-bai-fu/Fu-zi | Aconitum coreanum/ Aconitum carmichaelii | 155880/157040 | KU318669/KY407560 | Park et al. (2017)                           |
| Ye-Xing-Ba   | Scrophularia dentata          | 152553      | KT428154      | Né et al. (2016)                             |
| Bei Mu       | Fritillaria thunbergii        | 152155      | KY646165      | Moon et al. (2018)                           |
| Di Huang     | Rehmannia chingii,            | 154055      | KX426347      | Zeng et al. (2016)                           |
| San Qi       | Panax notoginseng             | 156324      | KT001509      | Zhang et al. (2016)                          |
| Ginseng      | Panax ginseng                 | 156356      | KM067388      | Kim et al. (2015)                            |
| Rough hedge parsley | Taraxacum officinarum   | 157855      | MN105615      | Yao et al. (2019)                            |
| Chervil      | Anthriscus cerefolium         | 154719      | NC_015113     | Downie and Jansen (2015)                      |
Table 2 Continued

| Common name | Species               | Genome size | Accession No. | References |
|-------------|----------------------|-------------|---------------|------------|
| Ma Huang    | Ephedra equisetina   | 109558      | MH161420      | Chen et al. (2019) |
|             | Ephedra intermedia   | 109667      | MH161421      |            |
|             | Ephedra sinica       | 109550      | MH161422      |            |
| Ginger      | Zingiber officinale  | 162621      | MH161428.     | Cui et al. (2019) |
| Bush clover | Lespedeza cuneata    | 149010      | MN268503      | Somaratne et al. (2019) |
| Goji berry  | Lycium chinense      | 155756      | MK040922      | Yang et al. (2019c) |
| Dwarf cardamom | Anomum compactum    | 163553      | MG000589      | Wu et al. (2018) |

*None* in the last column means the chloroplast genome sequences of species are available in NCBI with accession number but there is no publication.

2014; Stull et al., 2015) continue to be instructive about the evolution of the genome itself. Despite long-standing debate about if, when and how to combine data from different sources (molecules and morphology) or different genomes (nuclear and organellar), combining all plastid genes in a single analysis, perhaps with different evolutionary parameter values, has generally received consensus, as these genes are linked and represent a single chromosome (perhaps viewed as a single character, Doyle, 1992), and most studies have indeed combined plastid genes into a single analysis. Recent evidence, however, has revealed extensive topological discordance among trees built from individual plastid genes, both across angiosperms as a whole (Walker et al., 2019) and within legumes (Zhang et al., 2020). Although differences among plastid gene trees may arise due to biological causes—for example heteroplasmic recombination and gene transfer among plastid, nuclear and mitochondrial genomes—strong conflict is unexpected, particularly given that heteroplasm is considered rare (Sancho et al., 2018; Sullivan et al., 2017), limiting the possibility of recombination. The factors responsible for observed discordant plastid gene trees are not well understood on an empirical level, and more research is needed into heteroplasy and intergenomic transfer, although results to date point to both stochastic and systematic error, the latter arising due to misspecifications of the evolutionary model used in the phylogenetic analysis (Walker et al., 2019). Regardless of the cause of gene tree conflict, plastid genes, if combined into a single analysis, may confute multiple phylogenetic signals, muddying the overall inference of both topology and branch lengths, with consequences for downstream analyses of divergence times, diversification, character evolution and more; thus, greater exploration of plastid gene trees is needed in future studies.

Conception and advancement of chloroplast genetic engineering

At the dawn of the GMO revolution in the 1980s, introducing herbicide resistance genes into the nuclear genome, using Agrobacterium-mediated transformation (Shah et al., 1986) with the possibility of escape of transgenes via pollen, was publicly debated. Another major concern was the development of resistance in insects to bacterial insecticidal (Bt) proteins produced in plants at low expression levels via the nuclear genome. Therefore, the concept of chloroplast genetic engineering was first demonstrated in isolated chloroplasts (Daniell et al., 1991; Daniell et al., 1984; Daniell and McFadden, 1987; Daniell et al., 1983; Rebeiz et al., 1984), with the goal of reintroduction of chloroplasts into protoplasts to regenerate transplastomic lines (Bonnett, 1976; Bonnett and Eriksson, 1974) (Table 1). However, invention of the gene gun by John Sanford at Cornell University eliminated the need for the latter step and facilitated introduction of foreign DNA directly into chloroplasts of plant cells (Daniell, 1993; Daniell et al., 1984; Daniell et al., 1990; Ye et al., 1990) (Table 1). After 30 years of research on foreign gene expression in chloroplasts, >75% of foreign genes (Daniell et al., 2016a; Daniell et al., 2016b) use the psbA regulatory sequences used in this first report (Daniell et al., 1990). Although several methods for DNA delivery into chloroplasts have been reported, gene gun delivery is still the only reproducible method (Table 1). Goldschmidt-Clermont (1991) introduced the first selectable marker gene for chloroplasts—the aadA gene (Table 1), which was subsequently used for transforming the tobacco chloroplast genome (Svab and Maliga, 1993) and most other crops (Jin and Daniell, 2015). Although foreign genes were introduced into the transcriptionally silent spacer region (Svab and Maliga, 1993), transcriptionally active spacer region in the inverted repeat region of the chloroplast genome was proposed by Daniell et al (1998) and is now the most widely used spacer region for transgene integration (Daniell et al., 2016a, b; Daniell et al., 1998; Jin and Daniell, 2015).

Codon optimizer algorithm for expression of large human genes in chloroplasts

Chloroplast genomes have been engineered to express foreign genes from bacterial, fungal, protozoan or human genomes for various biotechnology applications. It is quite astounding that a bacterial operon from Bacillus thuringiensis was expressed utilizing bacterial regulatory sequences and achieved the highest level of Bt protein reported (DeCosa et al., 2001). Small human genes (insulin) are expressed at very high levels (up to 70% of leaf total protein (Boyhan and Daniell, 2011; Ruhlman et al., 2010)). However, when large human genes are introduced into the chloroplast genome, there are poorly translated. Therefore, a new codon optimization program was developed based on codon usage of the most highly expressed psbA genes from 133 sequenced chloroplast genomes (Kwon et al., 2016). The first iteration used a new codon optimizer algorithm that followed the codon usage hierarchy observed among sequenced psbA genes. Synonymous codons for each amino acid were ranked according to their frequency of use. Codon optimization of the human blood clotting factor FVIII-HC (2262 bp) modified 406 codons out of 754 amino acids. Native FVIII-HC used the CTC leucine codon 11 times but the codon-optimized HC eliminated all CTC codons. Likewise, another rare codon, TCA (serine), used 16 times in the native FVIII-HC coding sequence, was completely eliminated after codon optimization. Collectively, the codon-optimization algorithm eliminated 105 rare codons from the FVIII-HC native human
gene, resulting in enhanced expression. Similar codon optimization of the CTB-FVIII-LC gene resulted in successful expression of the largest human blood protein (FVIII, 185 kDa monomer, with pentamer assembly) in chloroplasts (Kwon et al., 2018), and this has now advanced to commercial production and clinical trials. Codon-optimized Ace2 is now tested in the clinic to treat COVID-19 patients (Daniell, 2020).

From conception to commercialization of leaf enzymes

The global enzyme market was valued at US $11 billion in 2018; it is used in the food/beverage (37%), cleaning (26%), bioenergy (15%) and animal feed (13%) industries. Rising demand in detergents, textile and wastewater treatment applications are propelling significant growth in the enzyme industry. Enzymes are now used as biofertilizers and soil enhancers to accelerate growth through enhanced beneficial microbes. Carbohydrase led the product segment, with 48% market share in 2018. Globally, the Asia Pacific region leads the industrial enzyme demand ($4.6 billion), followed by Europe ($2.68 billion) and the United States has the largest (38%) market share of industrial enzymes.

Waxes and pectins decrease water absorption of native cotton fibres; contrary to common understanding, natural cotton fibre does not absorb water or dyes (Lin and Hsieh, 2001). To enhance absorption of harsh chemicals, alkaline pH and other severe conditions have been used, but these are environmentally problematic in that they release toxic effluents. Therefore, pectinases are used for textile bioscouring to enhance water absorbency of cotton fibres, without causing cellulose destruction. Likewise, lipases are used in the detergent industry to remove oils or butter stains (Jaeger and Reetz, 1998). However, performance of current commercial lipases is limited by minimal activities in alkaline pH or at higher temperature (detergents are alkaline and washing machines often use hot water).

Mannans are found in chocolate, ice cream and tomato ketchup, which are unfortunately the most common cloth stains (Chauhan et al., 2012). Mannanase hydrolyses insoluble mannann into water-soluble smaller oligosaccharides (Dhawan and Kaur, 2007), thereby facilitating stain removal. However, current mannanase products are not active through a wide range of temperatures or pH (Sarmiento et al., 2015). Mannannase is also used in the paper and pulp industries, bioethanol production, oil and gas well stimulation, food and feed, nutraceutical and pharmaceutical industries (Srivastava and Kapoor, 2017; Zyl et al., 2010). Endoglucanases and exoglucanases are useful in enhancing colour brightness and fabric softening (Agrawal, 2017). These enzymes are used in denim biowashing, biostoning and biopolishing to remove dye from the fibril surface with minimal damage to fabrics (Anish et al., 2007; Araújo et al., 2009; Miettinen-Oinonen and Suominen, 2002). Cocktails of cellulases, hemicellulases and pectinases are used for fruit juice clarification, concentration and reducing viscosity (Brito and Vaillant, 2012; Sharma et al., 2016).

Currently, industrial enzymes are produced in microbial systems that are expensive and that rely on decades-old processes. Building fermentation facilities and their maintenance free of contamination is the first challenge. Purification of enzymes from host cells and their formulation to increase concentration, stability and cold storage/transportation are all very expensive. Therefore, several enzyme products made in leaves and expressed in chloroplasts (Agrawal et al., 2011; Jin et al., 2011; Leelavathi et al., 2003; Verma et al., 2010) were launched recently by PhyloZyme in order to replace prohibitively expensive microbial enzyme processes (Daniell et al., 2019a; Daniell et al., 2019c; Kumari et al., 2019). This is the first report of commercially launched protein products made in leaves (Table 1, Figure 1). Leaf pectinases, endoglucanases, lipases or mannanases were validated with 23 commercial microbial enzyme products for textile (dye binding, removal, depilling), detergent (stain removal) or juice industry (clarification) applications. Most leaf enzymes function in broad pH/temperature ranges as crude leaf extracts without the need for purification, and leaf enzymes could be stored as lyophilized plant cells at ambient temperature for several months/years, without loss of enzyme activity.

CO₂ emission by fermentation and capture from chloroplasts

Commercially produced enzymes are widely used in daily life from coffee/juice in the morning, textiles worn during the day (enhancing water absorption, dye binding), cleaning of dishes/clothes (detergents), evening beverages (wine, beer) or digesting food (fat, carbohydrates). Currently, these enzymes are produced in yeast, fungi or bacteria in fermentation systems. Life cycle analysis of β-glucosidase, used for digesting milk products, revealed that, for each kg of β-glucosidase enzyme, 52 kg of CO₂ are released from fermentation, along with 18,140 kg wastewater and >15 kg of solid debris and carbon monoxide, methane and other toxic gases, in addition to environmental stress caused by toxic chemicals used in cleaning the fermenters used in enzyme production (Feijoo et al., 2017). Considering the millions of tons of microbial biomass produced every day, dealing with global CO₂ emission and toxic effluents is a major environmental challenge.

Here, we use cellulase production with the tobacco enzyme production platform to evaluate the carbon capturing potential of leaf enzyme technology. Assuming a standard tobacco production for 12 weeks, reaching a final leaf area index (LAI) of 2.5, and average photosynthetic properties as reported in (Schmidt et al., 2019), respiration of 30% of the total photosynthetic CO₂ uptake (Amthor, 1989; Amthor, 2000), then the total CO₂ uptake by such a tobacco canopy to produce 1 kg cellulase will be about 570 kg. If at the end of the growth cycle, if LAI is 6 instead, a net photosynthetic CO₂ uptake of 484 kg is predicted during the production of 1 kg cellulase. Replacement of enzymes made by fermentation by leaf enzymes could capture up to 570 kg CO₂ in addition to preventing 110 kg CO₂ release for a net capture of 680 kg CO₂ for each kg of enzyme/protein produced, altogether leading to decreased anthropogenic greenhouse gas emissions.

Affordable protein drugs and vaccines—from laboratory to the clinic

Although protein drugs like insulin save millions of lives, they are not affordable for a large majority of the global population. Indeed, insulin pumps cost $6,000-$12,000 but one third of the global population earn <$2 per day. Therefore, new approaches are needed to produce and deliver protein drugs (PDs) more cost effectively. Edible plant cells offer the opportunity to express and orally deliver PDs. Upon oral delivery, PDs are protected in the stomach from digestion by enzymes because human or animal enzymes do not break beta linkage of plant cell wall polymers. However, when intact plant cells reach the gut, enzymes released by commensal microbes in the gut digest plant cell walls, releasing PDs into the gut lumen (Kumar et al., 2020). When tags are fused to PDs, they cross gut epithelium and reach the circulatory or immune system (Figure 2). Therefore, oral delivery of PDs
Figure 1  Comparison of chloroplast and microbial enzymes on dye or stain removal, water absorption or juice clarification: (a) Endoglucanase expressed in chloroplasts (a3) removed dye from denim fabric uniformly unlike uneven removal using microbial enzyme (a2) when compared to untreated control (a1). (b) Chloroplast pectinase (b2, b3) clarified orange juice pulp similar to microbial pectinase (b4) when compared to untreated control (b1). (c) Chloroplast pectinase (c4) increased water absorption of cotton fabric similar to microbial enzyme (c3), when compared to controls with water (c1) or untransformed plant extract (c2). (d) Chloroplast mannanase (d4) removed cholate stain more efficiently than microbial enzyme (d3) when compared to untreated control (d1) or with detergent alone (d2). Antibiotic resistance gene was removed from lettuce transplastomic lines but not from tobacco. For further details, see Daniell et al, 2019 or Kumari et al, 2019.

Figure 2  Oral delivery of protein drugs made in lettuce chloroplasts: (a) Production of Ace2/Ang1-7 in lettuce chloroplasts, freeze drying and oral delivery of capsules. (b) Protection of intact plant cells expressing GFP from stomach acids and enzymes and delivery to the gut. Gut microbes release enzymes to digest the plant cell wall and release GFP to gut lumen, which is taken up by gut epithelial cells. (c) SARS-CoV-2 binds to the human Ace2 receptor via the spike protein (target of several vaccines), lowers Ace 2 levels and causes lung/heart injury through inflammation by deregulation of the RAS pathway. (d) Oral delivery of angiotensin-converting enzyme 2 and angiotensin-(1-7) bioencapsulated in plant cells attenuates pulmonary hypertension and reduces right ventricular systolic pressure, right ventricular hypertrophy. Fibrosis and pulmonary vessel wall thickness, which are the symptoms observed in COVID-19 patients. For further details, see Daniell (2020).
expressed in lettuce chloroplasts are developed to treat infectious or inherited diseases in several animal models and are now advanced to the clinic.

Recent advances include expression of PDs in marker-free chloroplast genomes by removal of the antibiotic resistance gene (Park et al., 2020; Daniell et al. 2020), production of lettuce expressing PDs in cGMP facilities (Daniell et al., 2019a, 2020; Park et al., 2020), evaluation of drug substance by toxicology and pharmacokinetic studies (Daniell et al., 2020) and recent FDA approval of PDs orally delivered in peanut cells to treat allergy (Tilles and Petroni, 2018; VICKENY et al., 2018). Two major examples are discussed below in the context of developing SARS-CoV-2 vaccine or treating acute/lethal lung/heart failure of COVID-19 patients (Figure 2).

There is an urgent need to develop new vaccine strategies for SARS-CoV-2. There are > 50 vaccine clinical trials currently in progress, most of them using the same antigen (spike protein) but produced using DNA, RNA or recombinant proteins (Amanat and Krammer, 2020). Because all of these are injectable vaccines (except Vaxart), they will primarily produce systemic immunity with suboptimal induction of mucosal surface immunity required to protect at viral entry points (Sui and Berzofsky, 2020). Induction of mucosal immunity will likely require oral vaccination or antigen delivery via other mucosal routes. In addition, the induction of suboptimal and/or short-lived responses (as seen already in repeat COVID-19 infection) including in the elderly would require additional boosters. In this context, recent development of a polio oral booster vaccine using lettuce chloroplasts, in collaboration with FDA and CDC laboratories, funded by the Gates foundation is highly significant. Oral delivery of codon-optimized polio viral protein 1 common to all polio viral serotypes generated both IgA and IgG1 specific antibodies and conferred protection against all three polio viral serotypes (Chan et al., 2016; Xiao and Daniell, 2017; Daniell et al, 2019b). While a single-dose injectable polio vaccine (IPV) did not produce significant levels of IgA and IgG1 titres decreased quickly, oral boosters with VP1 expressed in chloroplasts maintained both IgA and IgG1 up to 400 days. In case of Sabin 1 and Sabin 2, a single dose of IPV resulted in poor seropositivity (~20% or < 40%), but showed ~ 70% seropositivity against Sabin 3. Plant cell boosting showed higher seropositivity against Sabin 1, Sabin 2 and Sabin 3 (80-100% protection). When OPV2 was withdrawn by the World Health Organization because it could revert to virulence by point mutations or recombination with other enteroviruses, most developing countries could not afford a second dose of IPV. In a very similar situation, an affordable oral booster strategy is now developed for SARS-CoV-2 oral booster vaccination, especially to develop prolonged mucosal immunity.

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus responsible for coronavirus disease 2019 (COVID-19), infects host cells via the angiotensin-converting enzyme 2 (ACE2) receptor, resulting in vasoconstriction, hypercoagulability, myocardial/ischemic injury, fibrosis, inflammation associated with ACE2 down-regulation and/or cytokine storm (Du et al., 2020; Zoufaly et al., 2020). In healthy people, ACE2 is expressed in type II alveolar lung epithelial cells that produce surfactants and this protects alveoli from collapsing. Most importantly, ACE2 produces anti-inflammatory, cytoprotective angiotensin 1-7 (Ang 1-7) peptide via cleavage of the vasoconstrictor angiotensin II (AngII). Of potential therapeutic relevance, oral delivery of both ACE2 and Ang1-7 biencapsulated in plant cells significantly decreased right ventricular systolic pressure and improved pulmonary blood flow and right ventricle function in diseased hypertensive animals (Daniell, 2020; Shenoy et al., 2014, Figure 2). Conceivably, therapeutic delivery of ACE2/Ang1-7 could also be employed to restore a more favourable balance of Ang II and Ang 1-7 in patients with COVID-19 disease. In contrast to exogenously delivered truncated (transmembrane deleted) soluble ACE2 (Zoufaly et al., 2020), full-length ACE2 accumulates in the lungs at 10-fold higher concentrations than in the plasma, with no evidence of toxicity (Daniell et al., 2020). Based on these observations, evaluation of therapeutic efficacy and safety of supplementing ACE2 and Ang 1-7 with this existing product in non-critically ill COVID-19 patients in the hospital and at home is in progress. Investigational New Drug application has been filed and has gone through two rounds of FDA review, by employing an integrated Phases 1 and 2 clinical trial design. This clinical trial is planned to treat a growing number of COVID-19 patients to protect them from lung/extrapulmonary injury and heart failure.

**Future perspectives**

Sequenced chloroplast genomes are available for fewer than 70 genera of the ~ 3,000 species of cultivated crops and of fewer than 2,000 of the ~ 350,000 species of flowering plants; thus, further efforts are needed to increase the number of sequenced plastomes. Approximately half of every chloroplast genome contains intergenic spacer regions and regulatory sequences that are not conserved, but essential for chloroplast genetic engineering. The lack of intergenic spacer sequences is an important unmet need. In addition, several attempts to transform cereal chloroplast genomes have been unsuccessful, and advances in this area are needed to apply this concept to confer desired agronomic traits. Almost all successful chloroplast genetic engineering reported here used the gene gun and the aadA gene. Therefore, future investigations could focus on development of new selectable markers and DNA delivery methods to achieve plastid transformation in cereal crops.

Plastid DNA sequences will continue to be an important tool in elucidating relationships and evolutionary processes across not only green plants, but also other photosynthetic lineages of life. Of particular interest in the next decade will be more studies that compare patterns of nuclear and plastid phylogenetic relationships. It is anticipated that many more discordant patterns of relationships will be uncovered across the photosynthetic tree of life, providing novel insights into a deep and recurring history of reticulate evolution. Moreover, understanding the nature of the discordance in the phylogenetic signal of different plastome-encoded genes will yield further insight into the evolutionary constraints on the plastome and thus the utility of different regions for bioengineering.

Chloroplasts play important roles in sustaining life on this planet by providing carbohydrates, amino acids and lipids as sources of animal (including human) nutrition and O₂, which supports animal life and moderates global warming by trapping CO₂. In the past two decades, chloroplasts have been called upon to produce proteins (enzymes) used in everyday life and to reduce the cost of drugs to offer affordable health care. Major advances have been made in this area in the past few years. When clinical trials currently in progress are successful, an affordable protein drug production and delivery will be launched soon using chloroplasts as bioreactors. Recent advances augur well for chloroplasts to play a central role in both medicine and in sustaining life on Earth, replacing decades-old prohibitively expensive fermentation and drug delivery technologies.
Chloroplast genome engineering and phylogeny

Acknowledgements
Most of the recent publications on chloroplast biotechnology in the Daniell laboratory reviewed here were supported by funding from NIH grant R01 HL 107904, R01 HL 109442 and R01 HL 133191.

Conflicts of interest
Henry Daniell discloses conflict of interest due to several patents on chloroplast genetic engineering and technology founder of PhylloZyme that launched chloroplast enzyme products and biopharmaceuticals research supported in the past by Bayer, Novo Nordisk, Johnson & Johnson and currently by Shire/Takeda.

Author contributions
HD conceived this topic and wrote this review. SJ contributed to sections on crop chloroplast genomes and contents of tables. X-GZ contributed the section on CO2 emission by fermentation and capture by chloroplasts. MAG, DES and PSS contributed sections on chloroplast genome and phylogenetics.

References
Agrawal, B.J. (2017) Bio-stoning of denim- an environmental-friendly approach. Curr. Trends Biomed. Eng. Biosci. 1, 1–3.
Agrawal, P., Verma, D. and Daniell, H. (2011) Expression of Trichoderma reesei β-mannanase in tobacco chloroplasts and its utilization in lignocellulosic woody biomass hydrolysis. PLoS One, 6, e29302.
Amanat, F. and Krammer, F. (2020) SARS-CoV-2 Vaccines: Status Report. Immunity, 52, 583–589.
Amthor, J.S. (1989) Respiration and crop productivity. Q. Rev. Biol. 10, 271–273.
Amthor, J.S. (2000) The McCree-de wt-penning de vries-thornley respiration paradigms: 30 years later. Ann. Bot. 1–20.
Anish, R., Rahman, M.S. and Rao, M. (2007) Application of cellulases from an alkalothermophilic Thermomonospora sp. in biopolishing of denims. Biotechnol. Bioeng. 96, 48–56.
Apel, W. and Bock, R. (2009) Enhancement of carotenoid biosynthesis in transplastomic tomatoes by induced lycopene-to-provitamin A conversion. Plant Physiol. 151, 59–66.
Araujo, R., Casal, M. and Cavaco-Paulo, A. (2009) Application of enzymes for textile fibres processing. Biocatalysis Biotransformation, 26, 332–349.
Arlén, P.A., Falconer, R., Cherukumilli, S., Cole, A., Cole, A.M., Oishi, K.K. and Daniell, H. (2007) Field production and functional evaluation of chloroplast-derived interferon-alpha2b. Plant Biotechnol. J. 5, 511–525.
Arlén, P.A., Singleton, M., Adamovicz, J.J., Ding, Y., Davoodi-Semirani, A. and Daniell, H. (2008) Effective plague vaccination via oral delivery of plant cells expressing F1-V antigens in chloroplasts. Infect. Immun. 76, 3640–3650.
Atheyon, R.A., McComish, B.J., Shepherd, L.D., Berry, L.A., Albert, N.W. and Lockhart, P.J. (2010) Whole genome sequencing of enriched chloroplast DNA using the Illumina GAII platform. Plant Methods, 6, 22.
Banerjee, A. and Stefanovic, S. (2019) Caught in action: fine-scale plastome evolution in the parasitic plants of Cuscusa section Cerasthorae (Convolvulaceae). Plant Mol. Biol. 100, 621–634.
Baur, E. (1909) Das Wesen und die Erblichkeitsverhältnisse der “Varietates albomarginatae Hort” von Pelargonium zonale. Zschr indukt. Abst. u. Vererbungslehre, 1, 330–351.
Bausher, M.G., Singh, N.D., Lee, S.B., Jansen, R.K. and Daniell, H. (2006) The complete chloroplast genome sequence of Citrus sinensis (L) Osbeck var ‘Ridge Pineapple’: organization and phylogenetic relationships to other angiosperms. BMC Plant Biol. 6, 21.
Beanard, G., Hernández, P., Khadari, B., Dorado, G. and Savolainen, V. (2011) Genomic profiling of plastid DNA variation in the Mediterranean olive tree. BMC Plant Biol. 11, 80.
Bonnett, H.T. (1976) On the mechanism of the uptake of Vaucheria chloroplasts by carrot protoplasts treated with polyethylene glycol. Planta, 131, 229–233.
Bonnett, H.T. and Eriksson, T. (1974) Transfer of algal chloroplasts into protoplasts of higher plants. Planta, 120, 71–79.
Boyhan, D. and Daniel, H. (2011) Low-cost production of proinsulin in tobacco and lettuce chloroplasts for injectable or oral delivery of functional insulin and C-peptide. Plant Biotechnol. J. 9, 585–598.
Brito, B. and Vaillant, F. (2012) Enzymatic liquefaction of wall-cells from kent and tomato atkins mango fruits. Int. J. Food Sci. Nutr. Eng. 2, 76–84.
Brunsfeld, S.J., Solits, D.E. and Solits, P.S. (2001) Comparative phylogeography of northwestern North America: a synthesis. In Silverton, J. and Antonovics, J., eds. Integrating Ecology and Evolution in a Spatial Context (pp. 319–340). Cambridge, UK: Cambridge University Press.
Cao, J., Jiang, D., Zhao, Z., Yuan, S., Zhang, Y., Zhang, T., Zhong, W. et al. (2018) Development of Chloroplast Genomic Resources in Chinese Yam (Dioscorea polystachya). Biomed. Res. Int. 2018, 6293847.
Cavalier-Smith, T. and Lee, J.J. (1985) Protozoa as hosts for endosymbiosis and the conversion of symbionts into organelles. J. Protozool. 32, 376–379.
Chakraborti, S.K., Lutz, K.A., Lertherwiaywong, B., Svab, Z. and Maliga, P. (2006) Expression of the crynAAb 81 gene in tobacco chloroplasts confers resistance to potato tuber moth. Transgenic Res. 15, 481–488.
Chen, H.T., Xiao, Y., Weldon, W.C., Oberste, S.M., Chunakov, K. and Daniell, H. (2016) Cold chain and virus-free chloroplast-made booster vaccine to confer immunity against different poliovirus serotypes. Plant Biotechnol. J. 14, 2190–2200.
Chase, M.W., Solits, D.E., Omlstead, R.G., Morgan, D., Les, D.H., Mishler, B.D., Duvall, M.R. et al. (1993) Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene rbcL. Ann. Mo. Bot. Gard. 80, 528–580.
Chauhan, P.S., Puri, N., Sharma, P. and Gupta, N. (2012) Mannanases: microbial sources, production, properties and potential biotechnological applications. Appl. Microbiol. Biotechnol. 93, 1817–1830.
Chen, X., Cui, Y., Nie, L., Hu, H., Xu, Z., Sun, W., Gao, T. et al. (2019) Identification and phylogenetic analysis of the complete chloroplast genomes of three ephedra herbs containing ephedrine. Biomed. Res. Int. 2019, 1–10.
Chen, T., Hu, G.-F., Wang, Y., Chen, Q., Wang, L., Zhang, J., Tang, H.-R., et al. (2018) Characterization of complete chloroplast genome and phylogenetic analysis of sweet cherry Cerasus avium (L.) Moench (Prunoeideae, Rosaceae). Mitochondrial DNA Part B 3, 1274–1275.
Chung, H.-J., Jung, J.D., Park, H.W., Kim, J.H., Cha, H.H., Min, S.R., Jeong, W.J., et al. (2006) The complete chloroplast genome sequences of Solanum tuberosum and comparative analysis with Solanaceae species identified the presence of a 241-bp deletion in cultivated potato chloroplast DNA sequence. Plant Cell Rep. 25, 1369–1379.
Correns, C.E. (1909) Vererbungsversuche mit blass(gelb)grunen und bunblättrigen Sippen bei Mirabilis jalapa. Z. Indukt Abstamm Vererbungsl. 1, 291–329.
Cui, Y., Nie, L., Sun, W., Xu, Z., Wang, Y., Yu, J., Song, J., et al. (2019) Comparative and phylogenetic analyses of ginger (Zingiber officinale) in the family Zingiberales based on the complete chloroplast genome. Plants 8(8), 283.
Curci, P.L., De Paola, D., Danzi, D., Vendramin, G.G. and Sonante, G. (2015) Complete chloroplast genome of the multifunctional crop globe artichoke and comparison with other Asteraceae. PLoS One 10, e0120589.
Daniell, H. (1993) Foreign gene expression in chloroplasts of higher plants mediated by tungsten particle bombardment. Methods Enzymol. 217, 536–556.
Daniell, H. (2002) Molecular strategies for gene containment in transgenic crops. Nat. Biotechnol. 20, 581–586.
Daniell, H. (2007) Transgene containment by maternal inheritance: effective or elusive? Proc. Natl Acad. Sci. USA, 104, 6879–6880.
Daniell, H. (2020) From conception to COVID-19: an arduous journey of tribulations of racism and triumphs. Plant Biotechnol. J. 18, 2147–2154.
Daniell, H., Chan, H.T. and Pasoreck, E.K. (2016a) Vaccination via chloroplast genetics: affordable protein drugs for the prevention and treatment of inherited or infectious human diseases. Annu. Rev. Genet. 50, 595–618.

Daniell, H., Datta, R., Varma, S., Gray, S. and Lee, S.B. (1998) Containment of herbicide resistance through genetic engineering of the chloroplast genome. Nat. Biotechnol. 16, 345–348.

Daniell, H., Krishnan, M. and McFadden, B.F. (1991) Transient expression of β-glucuronidase in different cellular compartments following biolistic delivery of foreign DNA into wheat leaves and calli. Plant Cell Rep. 9, 615–619.

Daniell, H., Ranganathan, M. and Gnanam, A. (1984) Radiosiotopic evidence for the polypectides associated with photosystem II. Biochem. Biophys. Res. Commun. 125, 988–995.

Daniell, H., Liu Clarke, J. (2019) Production of tetravalent dengue virus (dengue virus 1-4) vaccines by oral administration of transgenic tobacco plants. Plant Biotechnol. J. 17, 134.

Daniell, H. and Liu Clarke, J. (2019) Production of tetravalent dengue virus (dengue virus 1-4) vaccines by oral administration of transgenic tobacco plants. Plant Biotechnol. J. 17, 134.

Daniell, H., Kristensen, M., Ranganathan, M. and Gnanam, A. (1984) Radiosiotopic evidence for the polypeptide associated with photosystem II. Biochem. Biophys. Res. Commun. 125, 988–995.

Daniell, H., Liu Clarke, J. (2019) Production of tetravalent dengue virus (dengue virus 1-4) vaccines by oral administration of transgenic tobacco plants. Plant Biotechnol. J. 17, 134.
Leelavathi, S., Gupta, N., Maiti, S., Ghosh, A. and Reddy, V.S. (2003) Overproduction of an algal- and thermo-stable xylanase in tobacco chloroplasts and efficient recovery of the enzyme. Mol. Breed. 11, 59–67.

Leff, J., Mandel, M., Epstein, H.T. and Schiff, J.A. (1963) DNA satellites from cells of green and aplastic algae. Biochem. Biophys. Res. Commun. 13, 126–130.

Li, H., Cao, H., Cai, Y.F., Wang, J.H., Qu, S.P. and Huang, X.Q. (2014) The complete chloroplast genome sequence of sugar beet (Beta vulgaris ssp. vulgaris). DNA Seq. 5, 9040.

Li, H.T., Yi, T.S., Gao, L.M., Ma, P.F., Zhang, T., Yang, J.B., Gitzendanner, M.A. et al. (2019) Origin of angiosperms and the puzzle of the Jurassic gap. Nat Plants 5, 461–470.

Li, P., Zhang, S., Li, F., Zhang, S., Zhang, H., Wang, X., Sun, R. et al. (2017) A Phylogenetic Analysis of Chloroplast Genomes Elucidates the Relationships of the Six Economically Important Brassica Species Comprising the Triangle of U. Front Plant Sci 8, 111.

Lin, C.H. and Hsieh, Y.L. (2001) Direct Scouring of Greige Cotton Fabrics with Trichoderma reesei endoglucanases and use of the new cellulase preparations in producing the stonewashed effect on denim fabric. Appl. Environ. Microbiol. 68, 3956–3964.

Meng, D.S., LiuZhenggang, X.Z., HuWenzhen, K.L. and Wu. (2018) The complete plastid genome of an economic plant. Camellia sinensis cultivar Anhuva. Mitochondrial DNA Part B: China.

Mereschkowsky, C. (1905) U’ ber Natur und Ursprung der Chromatoporen im Pflanzenreiche. Biologisches Centralblatt 25, 593–604.

Middleton, C.P., Senerchia, N., Stein, N., Akhunov, E.D., Keller, B., Wicker, T. and Kilian, B. (2014) Sequencing of chloroplast genomes from wheat, barley, rye and their relatives provides a detailed insight into the evolution of the Triticeae tribe. PLoS One 9, e95761.

Mettlen-Donnen, A. and Suomalainen, P. (2002) Enhanced production of Trichoderma reesei endoglucanases and use of the new cellulase preparations in preparing the stonewashed effect on denim fabric. Plant Biotech. (N Y) 19, 1862–1875.

Moore, M.J., Bell, C.D., Solits, P.S. and Solits, D.E. (2007) Using plastid genome data to resolve enigmatic relationships among basal angiosperms. Proc. Natl Acad. Sci. USA 104, 19363–19368.
Vickery, C.R., Cardenas, J., Bowman, M.E., Burkart, M.D., Da Silva, N.A. and Noel, J.P. (2018) A coupled in vitro/in vivo approach for engineering a heterologous type III PKS to enhance polyketide biosynthesis in Saccharomyces cerevisiae. *Biotechnol. Bioeng.* 115, 1394–1402.

Viitanen, P.V., Devine, A.L., Khan, M.S., Deuel, D.L., Van Dyk, D.E. and Daniell, H. (2004) Metabolic engineering of the chloroplast genome using the Echerichia coli ubiC gene reveals that chorismate is a readily abundant plant precursor for p-hydroxybenzoic acid biosynthesis. *Plant Physiol.* 136, 4048–4060.

Walker, J.F., Walker-Hale, N., Vargas, O.M., Larson, D.A. and Stull, G.W. (2019) Characterizing gene tree conflict in plastome-inferred phylogenies. *Peer J.* 7, e7747.

Wang, S. and Gao, L.Z. (2015) The complete chloroplast genome of an irreplaceable dietary and model crop, foxtail millet (Setaria italica). *Mitochondrial DNA* 27, 1–2.

Wang, K., Li, L., Zhao, M., Li, S., Sun, H., Lv, Y. and Wang, Y. (2017) Characterization of the complete chloroplast genome of longan (Dimocarpus longan Lour.) using illumina paired-end sequencing. *Mitochondrial DNA Part B* 2, 904–906.

Wu, M.L., Li, Q., Xu, J. and Li, X.W. (2018) Complete chloroplast genome of the medicinal plant Amomum compactum: gene organization, comparative analysis, and phylogenetic relationships within Zingiberaceae. *Chin Med.* 13, 10.

Xiao, Y. and Daniell, H. (2017) Long-term evaluation of mucosal and systemic immunity and protection conferred by different polio booster vaccines. *Vaccine* 35, 5418–5425.

Xie, F., Lin, S., Lai, Q., Lian, H. and He, B. (2019) The complete plastid genome of Chinese cinnamon, Cinnamomum Aromaticum Nees (Lauraceae). *Mitochondrial DNA Part B* 4, 3831–3833.

Xu, F., He, L., Gao, S., Su, Y., Li, F. and Xu, L. (2019) Comparative Analysis of two Sugarcane Ancestors Saccharum officinarum and S. spontaneum based on Complete Chloroplast Genome Sequences and Photosynthetic Ability in Cold Stress. *International Journal of molecular sciences* 20, 3828.

Yan, L., Lai, X., Li, X., Wei, C., Tan, X. and Zhang, Y. (2015) Analyses of the complete genome and gene expression of chloroplasst of sweet potato (Ipomoea batata). *PLOS One* 10, e0124083.

Yan, M., Zhao, X., Zhou, J., Huo, Y., Ding, Y. and Yuan, Z. (2019) The complete chloroplast genome of cultivated apple (Malus domestica Cv. ‘Yantai Fuji’). *Mitochondrial DNA Part B* 4, 1213–1216.

Yang, B., Chen, Y. and Zhang, C. (2019a) Characterization of the complete chloroplast genome of the Chinese sorghum, Sorghum bicolor from China. *Mitochondrial DNA Part B* 4, 1421–1423.

Yang, Z., Huang, Y., An, W., Zheng, X., Huang, S. and Liang, L. (2019c) Sequencing and Structural Analysis of the Complete Chloroplast Genome of the Medicinal Plant Lycium chinense Mill. *Plants (Basel)* 87.

Yang, S., Sun, X., Wang, L., Jiang, X. and Zhong, Q. (2019b) The complete chloroplast genome sequence of date palm (Phoenix dactylifera L.). *PLOS One* 5, e12762.

Yao, X.Y., Chen, Z.X. and Wang, Q.Z. (2019) The complete chloroplast genome of Torilis scabra (Apiaceae). *Mitochondrial DNA Part B* 4, 2914–2915.

Yao, X., Tang, P., Li, Z., Li, D., Liu, Y. and Huang, H. (2015) The First Complete Chloroplast Genome Sequences in Actinidiaceae: Genome Structure and Comparative Analysis. *PLOS One* 10, e0129347.

Ye, G.N., Daniell, H. and Sanford, J.C. (1990) Optimization of delivery of foreign DNA into higher-plant chloroplasts. *Plant Mol. Biol.* 15, 809–819.

Ye, G.N., Hajdukiewicz, P.T., Broyles, D., Rodriguez, D., Xu, C.W., Nehra, N. and Staub, J.M. (2001) Plastid-expressed 5-enolpyruvylshikimate-3-phosphate synthase genes provide high level glyphosate tolerance in tobacco. *Plant J.* 25, 261–270.

Ye, C.Y., Lin, Z., Li, G., Wang, Y.Y., Qiu, J., Fu, F., Zhang, H. et al. (2014) Echinocloa chloroplast genomes: insights into the evolution and taxonomic identification of two weedy species. *PLOS One* 9, e113657.

Yi, D.K. and Kim, K.J. (2012) Complete chloroplast genome sequences of important oilseed crop Sesamum indicum *L.* *PLOS One* 7, e35872.

© 2021 The Authors. Plant Biotechnology Journal published by Society for Experimental Biology and The Association of Applied Biologists and John Wiley & Sons Ltd., 19, 430–447.
Yin, D., Wang, Y., Zhang, X., Ma, X., He, X. and Zhang, J. (2017) Development of chloroplast genome resources for peanut (Arachis hypogaea L.) and other species of Arachis. Sci Rep 7, 11649.

Yoo, B.C., Yadav, N.S., Orozco, E.M. Jr and Sakai, H. (2020) Cas9/gRNA-mediated genome editing of yeast mitochondria and Chlamydomonas chloroplasts. PeerJ 8, e8362.

Zeng, S., Zhao, J., Han, K. and Liu, Z.L. (2016) Complete chloroplast genome sequences of Rehmannia chinensis, an endemic and endangered herb. Conservation Genetics Resources 8, 1–3.

Zhang, R., Wang, Y.H., Jin, J.J., Stull, G.W., Bruneau, A., Cardoso, D., De Queiroz, L.P. et al. (2020) Exploration of Plastid Phylogenomic Conflict Yields New Insights into the Deep Relationships of Leguminosae. Syst. Biol. 69, 613–622.

Zhang, C., Zhu, Q., Liu, S., Gao, P., Zhu, Z., Wang, X. and Luan, F. (2018) The complete chloroplast genome sequence of the Cucurbita pepo L. (Cucurbitaceae). Mitochondrial DNA Part B 3, 717–718.

Zhao, Z., Gao, A., Huang, J. and Luo, R. (2019) The complete sequence of chloroplast genome from mango (Mangifera indica var GuiFei). Mitochondrial DNA Part B 4, 1916–1917.

Zhu, Q., Guang, Y., Xiao, X., Shan, N., Wan, C. and Yang, Y. (2019) The complete chloroplast genome sequence of the Apium graveolens L. (Apiaceae). Mitochondrial DNA Part B 4, 463–464.

Chloroplast genome engineering and phylogeny