Notes on Photosynthesis and Plant Biotechnology

Department of Life Sciences
Imperial College London

Molecular Biology of the Chloroplast – Prof. Peter Nixon

Introduction - Solar Energy Conversion

Energy demand driven by a rising global population must increasingly be satisfied from renewable alternatives to fossil fuels, as the latter release extensive amounts of greenhouse gases with devastating consequences for our ecosystem. Solar power is considered to be a particularly attractive source, as on average the Earth receives around 10,000 times more energy from the Sun in a given time than is required by human consumption. In particular, there is 14 TW energy gap predicted to arise by 2050. Solar energy is approximately 100000 TW every day. How can we capture it and use it as fuel? Exploit photosynthesis!
In addition, since the industrial revolution the carbon dioxide concentration in the atmosphere has almost double (now we are in the range of 400ppm), certainly as a result of human oxidative activities. How can we sequester it from the atmosphere and use it as feedstock to counteract our damages? Exploit photosynthesis!

Oxygenic photosynthesis is an energy transduction process that takes plants in plants, algae and cyanobacteria. In this process, solar energy is converted into chemical energy using the energy from the photons. Inorganic carbon is reduced using electrons extracted from water molecules (photolysis) and is ultimately fixed in organic carbon compounds (sugars, carbohydrates), necessary for the growth and biomass of the phototrophs. Oxygen is released as waste product of water oxidation. Evolution of oxygenic photosynthesis is thought to be the factor responsible for the oxygenation of the Earth. In the thylakoid membranes (found in chloroplasts in eukaryotes) the light reactions take place. Where light-harvesting chlorophyll molecules, photosystems are found and oxygen, ATP and NADPH are evolved. In the stroma (like cytoplasm) you can find a lot of Rubisco and is where the dark reactions take place (CO2 fixation, sugar and starch synthesis).

In C3 plants, 3ATP and 2NADPH molecules are needed to fix $1CO_2$ into triose phosphate in the Calvin Cycle. This means that the ATP/NADPH ratio required to drive carbon fixation is 1.5. How much ATP and NADPH is produced in the light reactions?
For every pair of electrons, 6 protons are pumped ($H^+/e = 6/2 = 3$)
The formation of NADPH requires a pair of electrons ($2e^-/NADPH = 6H^+/NADPH$).
14 H^+ are thought to be needed by ATP synthase to make 3 ATP (14/3 = 4.67 H^+/ATP).

Therefore the ATP synthesised per NADPH is $\frac{6 \frac{H^+}{NADPH}}{4.67 \frac{H^+}{ATP}}$ = 1.29 ATP/NADPH.

So apparently linear electron flow (LEF) does not produce enough ATP for the Calvin cycle.

It is now thought (even though only recently is being studied) that cyclic electron flow (CEF) makes up the shortfall. The pathways for cyclic electron flow around PSI are still not understood (it was first observed by Arnold and co-workers in 1954). What we know is that reduced ferredoxin, Fd, (which usually is reduced by FNR and itself reduces a lot of other enzymes involved in nitrogen, sulfur, carbon assimilation and many others), can catalyse a cycle around PSI. In this cycle (CEF) electrons are eventually returned to P700 via the cytochrome b6f complex. CEF is readily activated if electron flow from PSII is cut off, for example with the inhibitor DCMU. In *Arabidopsis* the CEF pathway involves the NAD(P)H dehydrogenase enzyme (NDH), which apparently uses reduced Fd to reduce plastoquinone, then re-oxidised by cytb6f. Another pathway involves the PGR5 protein. In the green alga Chlamydomonas reinhardtii, a protein super complex that catalyses CEF has been recently identified (Iwai et al., 2010).

An extensive membrane system is found within the chloroplast or in cyanobacteria., all the chlorophylls and other pigments are associated with these membrane, known as thylakoids (or sometimes lamellae). In higher plant chloroplasts, most of thylakoids are closely associated in stacks (known as grana /appressed membranes those not associated are known as stroma lamellae. The structure of these membranes is hard to study. Usually we thought of them with sort of fork modules. Now there is evidence that perhaps the grana form helical structures. (Pribil et al., 2014).

There are possible routes to enhance photosynthesis, which involve either modifying the light reactions, dark reactions or improving photoprotection. Or even doing artificial photosynthesis. These are discussed in detail in James Murray's notes (see below).
Another route to exploit photosynthesis different than increasing crop yield is to use algae as biofuel factories. These have several advantages; the main one being their photoautotrophic growth. They can sequester carbon dioxide while t the same time produce fuels only using sunlight energy and water as feedstock (and potentially wastewaters). They can grow on land and in water unsuitable for growth. They also have higher photosynthetic conversion efficiency than plants.
However, some challenges exist, such as the provision of water and nutrients (solvable using wastewater), the harvesting of biomass (solvable by expressing volatile compounds or using biofilm bioreactors), contamination of open ponds (solvable using biocontainment strategies) and the expenses of closed systems.
There was a lot of hype about algal biofuels however it was mainly a bubble. As low-value high-volume products are still not economically feasible to cover operational costs- Many see two stages of photobiomanufacturing. In the first stage phototrophs can be used to produce high value low-volume compounds, as these cover the expenses. As process efficiencies improve, production of cheaper molecules will be driven by large market sizes. Pamela Silver envisions that high-value cyanobacterial products will in turn facilitate the generation of products with progressively lower value, but larger markets (octane and sugar), by allowing the economical operation of large production facilities.

Hydrogen Bioproduction

Many microalgae have a H_2-centered metabolism in which hydrogen serves as a source of reductant. Interestingly, microalgae can link directly photosynthetic water oxidation to H_2 production by hydrogenases. H_2 is cool as its combustion does not release toxic products, but just water.

It was found by Gaffron and Rubin (1942) that *C. reinhardtii* is capable of producing hydrogen spikes after a period of anaerobic induction. Indeed, anaerobiosis is essential as hydrogenases are repressed by oxygen. Then it was observed that sulfur-deprived microalgae can give rise to sustained hydrogen production, up to 2mL hourly per liter (Mellis, 2010). The absence of sulfur prevents repair of photodamaged PSII. Thus, PSII stops evolving oxygen, and when the rate of oxygen photoproduced matches that consumed by the mitochondrial respiration the cell is in anaerobic condition, explaining why sulfur depletion leads to hydrogen production.

Three pathways link electron flux to the hydrogenase:

1. **Biophotolysis** (direct pathway, PSII-dependent): the electrons extracted from water by PSII go to ferredoxin (FDX), which usually reduces ferredoxin-NADP oxidoreductase (FNR). In anoxic situations, FCX can reduce the hydrogenase, catalysing the reversible reduction of protons into molecular hydrogen.
2. **Photofermentation** (PSII-independent): reductant released from glycolytic degradation of glucose are transferred directly to plastoquinone pool by the NADP/plastoquinone oxidoreductase (NPQR), bypassing PSII. After illumination, these end up in the classic photosynthetic electron chain, reducing plastocyanin, re-energised by PSI and end up in ferredoxin as in the direct pathway.
3. **Dark Fermentation**: under dark anoxia, the pyruvate-ferredoxin-oxidoreductase (PFR) transfers electrons from pyruvate to the hydrogenase via ferredoxin.

Rational Improvement Designs:

- Engineer more resistant (oxygen-insensitive Fe-Fe hydrogenases using directed evolution?)
- Increase flux of electrons going to hydrogenases
- Modulate oxygen levels in the cell (e.g. increase O2 consumption by manipulating mitochondrial metabolism)
- Improve bioreactor design
- Or through serendipity: Random mutagenesis/screening identified *moc1* mutant which has 10-fold better rate of hydrogen production (2% energy conversion efficiency) than WT. MOC1 is involved in regulating transcription termination in mitochondrion – indirect effect on redox state of chloroplast

Another approach is to exploit photosynthesis for electricity production (in BPVs). You can even 3D print cyanobacterial biofilms (Sawa et al., 2017).

Biological photovoltaics (BPVs) are emerging as an environmentally friendly and low-cost approach to harvest solar energy and convert it into electrical current. In phototrophic organisms, light is converted into high-energy charge-separated electron–hole pairs, and the excited electrons are transferred through a number of intracellular electron carriers, with a fraction eventually exported across the cell

membrane and released to the external environment. In BPVs, these secreted electrons are directed to an electrode (anode) and from there allowed to flow to a more positive potential electrode (cathode) through an external circuit, thus generating current. Simultaneously, the protons released by the cells diffuse from the anodic chamber to the cathodic one where water is re-formed on an appropriate catalyst. This process leads to the generation of current without release of any chemical side-products. A proton-permeable membrane separates the anodic chamber from the cathodic one, ensuring that electrons travel only via the external load.

Cyanobacteria as biorefineries

Cyanobacteria are photosynthetic prokaryotes that can use photon energy to ultimately transfer electrons from water to carbon dioxide, generating more reduced molecules in the process. The introduction of heterologous, mostly catabolic, pathways into the metabolism of cyanobacteria allows production of a wide range of fuel and commodity products from CO_2, light and water
Because of their high energy content, fatty acids and their alcohol derivatives and alkane derivatives are attractive fuels. Photosynthetic production and excretion of free fatty acids (FFAs), through overexpression of a thioesterase, was achieved in Synechocystis 6803 (Liu et al., 2011) and *Synechococcus* 7942 (Ruffing and Jones, 2012). Using an alternative approach, that is, overexpression of the endogenous acyl-ACP reductase, Kaiser et al (2013). demonstrated excretion of FFAs in *Synechococcus* 7942 . In the same study, production of triacylglycerols and wax esters was reported. Overproduction of alkane biosynthesis genes from various cyanobacteria in *Synechocystis* 6803 led to an increase of heptadecane and heptadecene content (Wang et al., 2013). Mutants harbouring NADPH-dependent fatty acyl-CoA reductase showed increased levels of C_{15}–C_{17} fatty alcohols (Yao et al., 2014).

For all these cyanobacteria applications, bioreactor design is very important as GMOs cannot be released in the environment. Biocontainment is a major obstacle. Selao et al. (submitted) reported the use of melamine as an alternative N source. GM cyanobacteria are deprived of the nitrate transporters etc… and can only break down melamine to get nitrogen (using a transgene). Recently a similar strategy has been published: cyanobacteria have been engineered to grow only on phosphite (Pt) and not phosphates for biocontainment (Motomura et al., 2018).

Plastid Evolution
Chloroplasts are semiautonomous cell organelles that contain DNA and ribosomes, which code for and make proteins involved in the photosynthetic apparatus.
Chloroplasts are the prototypical members of a diverse family of organelles, the plastids. In plants, other plastids are the amyloplasts (starch-rich in seeds, roots and tubers), chromoplasts (carotenoids pigments in flowers and fruits). These plastids all come from the undifferentiated protoplasts found in meristems and reproductive tissues.

However, not all the photosynthetic proteins come from the plastid genomes, some are encoded in the nuclear genome of the host eukaryotic cell.
Analysing this division of genetic labour between the two genomes reveals important insights about the evolutionary histories of chloroplasts, and their integration with their hosts.
The endosymbiosis theory describes the chloroplasts as organellar relics of a symbiotic association between a protoeukaryote and a bacterium.
This theory was proposed a century ago by Shimper (as an alternative to the autogenous theory of Cavalier-Smith), but it was thanks to the efforts of the evolutionary biologist Lynn Margulis in the 90s that it was widely accepted by the scientific community.

Evidence for endosymbiosis theory:

- Circular DNA genomes with no histones
- Bacterial-like 70S ribosomes
- Sensitivity patterns to protein synthesis inhibitors
- Protein translation start with N-formyl-methionine
- No poly-adenylation of chloroplast mRNA
- Bacterial-like promoters and RBSs (Shine-Dalgarno)

But probably what is most convincing is doing comparative genomics analysis between chloroplast DNA and that of cyanobacteria. Because of the high similarity, it has been suggested that the endosymbiont was a freshwater cyanobacterium (possibly even capable of Nitrogen fixation), and it was taken up by a heterotrophic protest approximately 1.5 billion years ago.

While symbiogenesis is now widely accepted, the question that remain is whether endosymbiosis happened once (monophyletic theory) or multiple times (polyphyletic theory) over the evolutionary history. The monophyletic theory states that a single primary endosymbiotic event took place, and all the chloroplasts are derived from that cell. The polyphyletic theory states that multiple endosymbiosis events took place.

It is widely assumed that the primary endosymbiosis was a unique event, although Howe et al. (2008) questioned the reliability of this interference There is recent evidence of an independent primary endosymbiotic acquisition of a cyanobacterium by the rhizarian amoeba *Paulinella chromatophora.* The fact that no known organism is configured exactly as required by the two theories, led Howe et al. (2008) to propose a compromise between the two.
The evidence now suggests that several groups of eukaryotic algae originated form secondary endosymbiosis, where a eukaryotic alga (result of a primary endosymbiosis event) was incorporated into a second host. Also it is possible that this secondary endosymbiont was engulfed by a third host, giving rise to tertiary plastids.
Higher plants chloroplasts (products of primary endosymbiosis) are surrounded by a double membrane, which is derived from the inner cell membrane of the original symbiotic cyanobacterium plus the membrane of the food vacuole produced by phagocytosis during the engulfment (or perhaps are derived from inner and outer membrane of the pioneer cyanobacterium).
The same applies for secondary endosymbionts, thus giving rise to 4 membranes.
Further evidence for secondary endosymbiosis is the presence of a vestigial nucleus, called nucleomorph, between the second and third membranes.

Sometimes over plastid evolutionary history, the ability to do photosynthesis has been lost but plastids did not disappear. This appears to be true for a number of parasitic organisms, including the malaria-carrier *Plasmodium falciparum* and toxoplasmosis-pathogen *Toxoplasma gondii*.
In some lineages, the entire cells of prokaryotic or eukaryotic photosynthetic symbionts are retained, referred to as 'photosymbionts'. In other cases, the host specifically harvests and preserves chloroplasts from photosynthetic prey, generating structures termed 'kleptoplasts'. It is interesting that, whereas so many eukaryotic lineages acquire photosymbionts or kleptoplasts, fewer have acquired permanent chloroplasts.

Plastid-Nucleus Communication

Biogenic control (nucleus to plastid)

The expression of plastic components depends on nucleus-encoded factors. Plastid DNA expression is initiated by the nucleus-encoded RNA polymerase (NEP, phage T7-type). Then, PEP becomes active. PEP is homologous to prokaryote RNAP and has five subunits. Four of them are encoded in the plastid genome. The sigma factors, which confer promoter specificities, are encoded in the nucleus. As a result, the nucleus has transcriptional control over the chloroplast. Proplastid-to-chloroplast transition involves the expression of plastid genes that encode components of PSI and PSII, driven by PEP. Nucleus-encoded subunits (like LHCs) assemble around these core components). Import of nucleus-encoded factors relies on the TIC-TOC translocation system. Interestingly, whereas mitochondrial translocation systems rely on proton motive force plus ATP, chloroplast TIC-TOCs are PMF independent and only need ATP. But recent evidence also suggests that there might be a vesicular route (endomembrane transport) to the chloroplast, especially for glycosylated proteins. Photosynthesis-related genes in the nucleus often contain G-box promoter elements and thus are photoactivated.
Chlorophyll pigments and carotenoids are synthesised in the chloroplast with the MEP pathway (even though eukaryotic carotenoids biosynthesis pathway exists in the cytosol of plant cells). These are produced using Glu-tRNA as starting material, like haem and other tetrapyrroles.

Operational control (plastid to nucleus)

Barley mutants with defects in plastid ribosomes, fail to synthesise nucleus-encoded plastid enzymes. This led to the notion of plastid retrograde signalling. Using this direction of information flow, chloroplasts can relay their status to the nucleus for fine-tuning purposes (operational control). For example, haem is synthetized in the chloroplast and is a positive signal that promotes the expression of nuclear genes encoding components of the photosynthetic machinery. On the other hand, failures in the final steps of chlorophyll synthesis in the chloroplast results in repression of those same nuclear genes. High light or loss of electron sinks (like in drought) promotes generation of ROS, which leads to the production of B-cyclocitral (oxidation product of beta-carotene) and other factors (like MEcPP). These then activate the expression of antioxidant and defence genes in the nucleus (like ascorbate peroxidase). Generally, the MEP pathway acts as a global stress sensor in the plastid.

Accumulation of Mg-Protoporphyrin in the chloroplast is both necessary and sufficient to regulate the expression of many nuclear genes encoding chloroplastic proteins associated with photosynthesis (Strand et al., 2003).

Why is the plastome retained? And not transferred completely to the nucleus?
Following endosymbiosis, many genes were transferred to the nucleus and proteins reimported into the chloroplasts. Chloroplastic genes are way less than bacterial. Most of the info needed to make the chloroplast is encoded in the nucleus. IN the genomes of green-algal derived chloroplasts there are approx. 60-120 protein coding genes; in those of red-algal lineage there are more (roughly 200). However, this is way less than the 3000 protein coding genes found in *Synechocystis*. Where did they go?
Most likely, some of them have been lost. Why keeping cell wall biosynthesis and phicobilisomes genes if you are a green-algal or plant chloroplast?!
The fact that photosynthetic eukaryotes have multiple chloroplast actually facilitate this transfer (and each chloroplast has multiple copies of its genome). If one gene stochastically ends up in the nucleus, at least there is a backup that allows genetic tinkering and prevents disruption of chloroplastic functions. After that, it is easy to see how the chloroplastic copy might become redundant and lost completely.

Some hypotheses:

- Hydrophobicity of proteins encoded by retained genes (however, what about nuclear-encoded LHCs….)

- To have redox control of gene expression (CORR): However, CORR does not readily explain the retention of a genome in plastids of organisms such as Plasmodium. Arguably, a plastid gene location might also be beneficial in allowing direct control of expression in response to other aspects of plastid physiology, not just redox poise [32], so we could extend the CORR hypothesis to ‘CO- location for Biochemical RegulAtion’ (COBRA). However, it is not obvious which Plasmodium plastid genes would require control in this way

- To have operon co-transcription and simultaneous co-assembly of protein complexes

- To protect the genes from ROS mutations (however the hydrogenosome is a weird organelle in some eukaryotes that produces hydrogen, it does not perform oxygen metabolism, yet it transferred all of its genes to the nucleus).

- To have sexual recombination of chloroplastic genes (at the end the chloroplast is a clonal, asexual entity and might accumulate deleterious mutations, even though there is HGT between chloroplasts…)

- Essential tRNA hypothesis: Although there is overwhelming evidence for the mass transfer of protein-coding genes from the plastid to the nucleus, there is no evidence as far as we can determine for the transfer of plastid genes encoding RNA products, whether for rRNA, tRNA, or RNA components of ribonucleoproteins such as the signal recognition particle (SRP). The same applies for mitochondria. Maybe prokaryotic RNAs are forbidden in nucleocystosol as they are degraded by nucleases? The key precursor in the biosynthetic pathway for tetrapyrroles such as haem and chlorophyll is d-amino- laevulinic acid (ALA). In plants and algae, ALA synthesis takes place in the plastid with the first steps involving activation of glutamate by its cognate tRNA followed by conversion of the glutamyl moiety to glutamate-1-semi- aldehyde and then to ALA. The tRNA is encoded by the plastid gene trnE, and is unique amongst all the plastid tRNAs in that it has a role in both tetrapyrrole biosynthesis and protein biosynthesis. As first pointed out by Chris Howe and Alison Smith in 1991, non- photosynthetic plants would therefore need to retain trnE, as well as the machinery for its transcription synthesize the haem component of mitochondrial cytochromes, P450 cytochromes and other essential oxidative enzymes. The tRNA-Glu is presumably indispensable for ALA formation, but why could it not be replaced by an imported cytosolic 80S-type tRNAGlu thereby making trnE redundant? Because of its role in ALA synthesis, the plastid tRNA has to interact with glutamyl-tRNA reductase, as well as glutamyl-tRNA synthetase and elongation factor EF-Tu. It seems unlikely that the cytosolic counterpart could easily replace this dual-function tRNA; a point that is illustrated by the finding that a single base change in trnE from Euglena results in the loss of haem biosynthesis but leaves protein synthesis unaffected.

- Limited transfer window (Barbrook et al., 2006): argues that species with a single plastid per cell (monoplastidic) experience less intercompartmental ptDNA transfer than those with many plastids per cell (polyplastidic) because they have fewer plastids to donate ptDNA, and lysis of the plastid would almost certainly result in death to the cell, unlike the case for polyplastidic taxa. Transgenic studies in tobacco (whose leaves contain hundreds of chloroplasts per cell) showed that marker DNA inserted into the plastome is transferred to the nucleus at a surprisingly high rate. Conversely, similar studies in Chlamydomonas suggested that transfer rates in this organism are at least several orders of magnitude lower. Second, analysis of plant nuclear genomes reveals extensive evidence of plastid DNA-derived sequences (NUPTs), suggesting the continuous influx of plastid DNA into plant genomes. By contrast, NUPTs are much rarer in the nuclear genomes of Chlamydomonas and Plasmodium.

In terms of communication, chloroplasts have been shown to communicate also between each other. Expression of Green Fluorescent Protein in chloroplasts reveals

the presence of stromules. Stromules are induced after a pathogen attack and act as docking pads for different chloroplasts.

Chloroplast Transformation

Key innovation that made organelle transformation possible was the invention of the gene gun, a devide that bombards living cells with accelerated DNA-covered microparticles (gold or platinum beads). This technology is called biolistic (biological + ballistic) transformation. First transformation of chloroplast was in *C. reinhardtii* (which has a big chloroplast that occupies half of the cell size). Other chloroplast transformation methods are possible, but are more labour intensive as they require for example cell-wall removal or incubation of protoplasts in polyethylene glycol (PEG) (this method then requires to regenerate the plants from wall-less protoplasts, which is quite finicky).

Stable transformation of chloroplasts requires:

1. Integration of the transforming DNA into the resident plastid DNA (via homologous recombination, indeed efficiency of targeting is positively correlated to the lengths of flanking regions)
2. Elimination of all untransformed copies of polyploid plastid genome

Typical workflow for a transplantomic plant (e.g. potato):

1. Preparation of leaves for biolistic bombardment
2. Exposure of bombarded leaf explants to spectinomycin-containing medium
3. Selection of primary transplantomic lines (green spots in white tissues)
4. Regeneration round under spectinomycin for homoplasmy
5. Additional regenerations from stem sections (this must be quick otherwise gene conversion can result in loss of transgene from the marker, and ofc check with PCR and southern blotting)
6. Growth of homoplasmic transplantomic plants

Advantages of chloroplast engineering (over nuclear):

- Maternal inheritance (uniparental or non-mendelian) minimises pollen-mediated dispersal of transgene
- Specific targeting of genes (homologous recombination)
- Higher copy number (up to 10,000 per cell)
- Selectable markers can be removed
- No gene silencing
- The gene product is retained in the plastid
- Potential use of operons

Arguably though, the most significant advantages of plastid engineering lie in its unique precision (because of transgene integration by homologous recombination) and the possibility of stacking multiple transgenes in operons for co-expression as polycistronic mRNAs, which currently is not feasible in the nuclear genome.

Disadvantages:

- limited post-translational modifications (NO GLYCOSYLATION) and
- not routine for major crop plants (e.g. cereals)

An exciting discovery is that plastid genome can also be shared horizontally between plants, as demonstrated by grafting experiments. This opens a bright future for transformation!

A common error in the analysis of transplastomic lines is to mistake promiscuous plastid DNA in the nucleus or the mitochondrion for heteroplasmy. Weak wild-type-like hybridization signals in DNA gel blot analyses or wild-type-like bands in PCR assays that persist over the regeneration rounds often come from plastid DNA transferred to the nuclear or mitochondrial genome. In these cases, homoplasmy can be verified by Southern blots with purified plastid DNA and/or by crosses and segregation assays that demonstrate a lack of phenotypic segregation in the next generation

Measuring the degree of gene containment

1. To the pollen

In most crops, the chloroplasts are maternally inherited. This offers an advantage for transplantomic plants, as the genetically modified material is not dispersed in the environment via the pollen (paternal vector). However, how true is that? What is the probability that plastid DNA can indeed be dispersed via the pollen?
Ralph Bock and his group (Ruf et al., 2007) answered this question by developing an experimental workflow for stringent selection of occasionally paternally inherited plastids.
They used male sterile tobacco plants as maternal recipient lines. The donor was a male transplantomic plant with transformed chloroplast carrying an antibiotic resistance protein (aadA) for selection and reporter protein (GFP) for visualization/quantification. After successful selection and isolation of homoplasmic transformed chloroplasts, this was used to generate paternal line. Large-scale genetic crosses between transformed paternal line (pollen donor) and sterile male plant (pollen recipient) they screened for green tissue sectors (GFP-expressing chloroplasts), surrounded by white tissues as maternally inherited chloroplasts should not be spectinomycin resistant. So if there is green, it means those cells are spectinomycin resistant and GFP comes from paternal transplantomic line only. However, it is important to put also GFP as just spectinomycin resistant can be evolved by random mutation in 16s rRNA.
Using this method, it has been calculated that the transmission rate of paternal ptDNA via pollen is in the order of **3 * 10^-6.** This indicates that transplantomic is an effective tool to degrease gene dispersal. However, in cases where pollen transmission must be prevented altogether, other containment methods need to be used/developed.

2. To the nucleus

Genetic material can be transferred from the chloroplast to the nucleus (as occurred during endosymbiotic evolution). However, it is probably still an ongoing process. To measure this transfer rate, Huang et al. (2003) introduced a nucleus-specific

neomycin phosphotransferase gene (neoSTLS2) in the chloroplast genome. The transfer of cpDNA into nucleus was measured by screening for kanamycin-resistant seedlings in the progeny between WT female and transplantomic (with neoSTLS2) pollen. The neoSTLS2 gene features a constitutive plant viral promoter, CaMV-35S, and a nuclear intron. It was designed to be functional only when transposed to the nucleus and to confer kanamycin resistance in young seedlings. An intron, from the potato nuclear ST-LS1 gene, was integrated within the reading frame to prevent synthesis of a functional protein in the chloroplast. Frequency of transfer was calculated as **1 transposition event per 16,000 grains of pollen.**

Tools for Transplantomics:

1. **Creating marker-free plants**

After successful selection of homoplasmic transformants, the selectable marker is not needed anymore. Actually this results in continuous expression of selection protein and thus unnecessary genetic/metabolic burden for the cell. In addition, the selectable marker is usually a protein conferring antibiotic resistance (usually to the chloroplast inhibitor spectinomycin) and thus it is not ideal to have antibiotic resistance plant growing in field conditions. For this reason, it is ideal to remove the selectable marker after successful transformation.
Marker gene excision can be achieved with the Cre-Lox recombination system, derived from the P1 phage. Cre recombinase is nuclear encoded but targeted to the chloroplast, where it binds the two cis-regulatory elements loxP (34 bp), which flank the 5' and 3' of the selectable marker (aadA). Cre recombinase is targeted to the nucleus by insertion of a N-terminal transit peptide derived from Rubisco small subunit. Homologous recombination between these inverted flanking repeats results in the excision of the selectable marker and in a marker free transformed plastid DNA.

2. **Inducible expression of foreign genes**

It would be ideal to control gene expression by precise and reversible induction.
One solution can be using chemical inducers that regulate RNA sensors and thus control the translation of specific mRNAs in plastids. Ralph Bock's group has demonstrated this inducible expression of GFP reporter in tobacco using a theo riboswitch (Verhounig et al., 2010). The aptamer domain binds the metabolite (theophylline) and this triggers conformational changes in the expression platform, which either permit or prevent gene expression.

3. **Improving expression of operons**

Plastid genes are usually part of operons and expressed as polycistronic elements (one promoter for all of them but individual RBSs). The stable polycistronic mRNA is then edited by removing the ends, trimming and usually cleaving it into monocistronic elements. This occurs by specific endonucleolytic cleavage. (One exception is the psbE operon (4 polypeptides for PSII), which is not processed and remains tetracistronic).

Most of the other genes, if they do not get processed (by having some mutations in intercistronic sequences) are often not translatable/functional. mRNA secondary structure formation has been implicated in impaired translatability of unprocessed polycistronic precursors. However, it is hard to predict the secondary structures of RNA and thus hard to design translatable operons for the expression of multiple transgenes. This is quite bad as simultaneous expression of multiple transgenes in polycistronic elements is heralded as one of the main selling point of chloroplast transformation.
Ralph Bock's group identified a small sequence element, referred to as an intercistronic expression element (IEE), that mediates the efficient intercistronic cleavage of polycistronic mRNAs into stable monocistronic transcripts (Zhou et al., 2007). This is an intercistronic cis-element derived from the intercistronic region between psbT and psbH of the psbB operon.
By expressing bicistronically yfp and nptII, separated by this sequence, they showed that, while this element is not required for processing downstream of the first cistron to occur, it is essential to confer mRNA stability and translation of the second cistron. The identified IEE is small enough to serve as a universal tool for stacking of foreign genes in operons, and thus will help to extend the range of applications of transplantomic technology.
Tobacco IEEs are recognised by HAT repeats of RNA-binding protein HCF107, which binds to the cis-regulatory elements and stabilise the transcript, thus making translation more efficient (Hammani et al., 2012). However, recently it has been reported that using tobacco IEE can destabilise expression of the native psbH gene, as the available HCF are sequestered by the heterologous IEEs. Thus finding novel IEEs is crucial.

Legen et al. (2018) – always from Ralph Bock's group - proposed the use of binding sites for PPR proteins as IEE- like elements. They have showed that by placing target sites for PPR10, HCF152, CCR2 and the binding site of a not yet identified protein, located upstream of the *rpl12* reading frame, as IEEs in a synthetic *neo-egfp* bicistron in tobacco chloroplasts, GFP accumulates to various levels. These results indicate that alternative sequences, to the already popular IEE, can be developed and used as modulators and enhancers of genes expression in polycistronic constructs.

On the other side, Macedo-Osorio et al. (2018) recently identified two novel IEEs from *C. reinhardtii* derived from the intercistronic regions between psbN-psbH and tscA-chlN capable of expressing bicistronically AadA and GFP.

Plants as Chemical Factories using Transplantomic

Edible Vaccine Photobiomanufacturing

An interesting example of the use of plants as sustainable chemical microfactories is the expression of tetanus toxin in tobacco chloroplast for manufacturing of edible vaccines. If you express the C fragment of the heavy chain of tetC this is immunogenic but not toxic. Interestingly, expression in E. coli has problems with codon usage and

in yeast there is a cryptic poly-A site within the gene so it is hard to express it in in those industrial workhorses. Why not in bananas? Then you give vaccine-banana to the people (all vaccines to the people…) and avoid high-costs and cold-chain (and fuck the pharma companies). First you need to assemble a construct: tetC gene flanked by 5' UTR and 3'UTR under the control of the Prrn promoter is the best. Then shoot the vector into the chloroplast using biolistic bombardment. If you managed to do it, confirm that the gene has been integrated by Southern blotting and detect the transcript and protein by SDS-PAGE/Immunoblot (I would just sequence and mass spec in one day…). If successful, confirm that vaccine works by feeding this transgenic tobacco (yes I know it's still not in bananas unfortunately) to some rats. If they don't die, you're good!
The take home message is that TetC expression in chloroplast provides a potential route towards the development of safe, plant-based edible vaccines against tetanus.

Conversion of Lycopene into B-carotene in Tomato
Apel and Bock (2009) introduced two lycopene cyclase (one from a bacterium and one from a plant) into the chromoplast genome of tomato (*Solanum lycopersicum),* one gene per chloroplast though still. Expression of the plant cyclase in the tomato fruit resulted in increased carotenoid biosynthesis and specifically conversion of the tomato antioxidant lycopene into the vitamin A precursor, beta-carotene.

Introducing Cyanobacterial Carboxysomes in Plant Chloroplasts
Long et l., (2018) recently reported the addition of 4 genes from the *Cyanobium* cyanobacterium in tobacco chloroplasts. Interestingly, this is the most minimal carboxysome construct ever reported (only Rubisco small and large subunit and 2 carboxysomal self-assembling proteins). They introduced them into a multigene operon with intercistronic elements and all the standard transplantomic package… This resulted in increased photosynthetic efficiency.

Introducing Synthetic Metabolic Pathways with Multigene Operons
Bock group (Lu et al., 2013) also engineered a synthetic multigene operon encoding three enzymes of the tocochromanol (vitamin E and humans cannot synthetize it) biosynthetic pathway and transformed tomato chromoplasts. Interestingly tocopherols arise from joining of shikimate aromatic amino acid pathway and isoprenoid biosynthetic pathway. Also, they introduced the intercistronic expression elements and this massively increased expression. After proof in tobacco, they expressed them in tomato chromoplasts. Induction was also increased after cold perios (induction by reactive oxygen species).

Bohmert-Tatarev et al., (2011) have also introduced bacterial genes from the polyhydroxybutyrate pathway and demonstrated efficient PHB bioplastic expression in tobacco plastids. Interestingly, the polymers turned out to be non-cytotoxic and can be readily isolated from the plant cells. This is pretty nuts, as it opens the possibility of plastic factories made out of plants (and not fossil fuel dependent!!).

More recently, 4 enzymes of the artemisin biosynthetic pathway was also transferred from *Artemisia annua* into the chloroplast of tobacco (Fuentes et al., 2016).
Artemisinin, a C_{15} isoprenoid (sesquiterpene) naturally produced in the wild plant *Artemisia annua* (sweet wormwood, native to temperate Asia), is the main

ingredient of artemisinin combination therapies (ACTs), currently the only effective cure of malaria.
In the first step, they inserted 4 genes of artemisinin biosynthetic pathway into the chloroplast. This resulted in expression of artemisinic acid, which can be extracted from the plants and convert into artemisinin by simple chemical reactions.
After testing different arrangements of the genes in the chloroplast, the plant line with highest levels of artemisinic acid was used to introduce a set of "accessory" genes into the **nuclear** DNA, using new method (COSTREL : new synbio method that combines chloroplast transformation with combinatorial nuclear transformation and large-scale metabolic screening of supertransformed plant lines) .These accessory genes are not strictly required to make the drug, but they help to regulate the process in a largely unknown manner. They generated hundreds of genetically modified plant lines, each with different combinations of accessory genes. After screening, the yield of artemisinic acid in the best line (AO3-CS180) increased 77-fold.

Also more recently, Bock's group demonstrated chloroplastic expression of the HIV entry inhibitor griffithsin, and showed that it can be stored stably in dried tobacco leaves (Hoelscher et al., 2018). Exciting for the possibility of plants as expression and delivery platforms for pharmaceuticals.

Crop Protection by RNAi in Chloroplasts

Chloroplast transformation also enables the use of RNAi in plant, by enabling to express dsRNA without it being cleaved by the host RNA interference mechanism (plant RNAi pathways cleaves foreign dsRNA into siRNA) but keeping it functional in the plastid (cyanobacteria lack the RNAi pathway). Zhang et al., (2015) introduced dsRNA against the beta-actin gene of the Potato Colorado Beetle into potato chloroplast and demonstrated crop protection without chemical insecticides !

Engineering Photosynthesis – Dr. James W. Murray

Introduction

If the global population keeps on growing at the rate at which is growing today, there will be approximately 9 billion people on Earth by 2050. For this reason, there are concerns regarding food security, as the projected crop yield by 2050 will not be enough to sustain the ever growing population. Specifically, yield need to be increased by 70% to meet the values required to meet the global demand. Food ultimately derives from photosynthetic carbon fixation by phototrophs, which use sunlight energy and water to fix inorganic carbon (CO_2) into organic reduced carbon compounds (which heterotrophs oxidise to extract energy used to fuel biochemical activities, hence life).
The crop yield potential is directly proportional to the planted land area, the probability of light incidence onto that area, the efficiency of conversion of solar energy into chemical energy, and coefficient of partitioning that chemical energy into edible biomass. In order to increase yield potential, we thus need to increase some of those parameters. With overpopulation at the doors, land area is expected to decrease in the future thus we cannot play with that parameter. The green revolution already

increased the frequency of light interception and partition efficiency, for example by developing strains with more fruits and less stems or with bigger leaves to capture more sunlight. The parameter that has been left almost unchanged is the photosynthetic efficiency, which is also relatively well conserved between different phototrophs. Here, I discuss different strategies to increase the value of photosynthetic efficiency.

The light reactions

Photosynthesis is the process in which light energy is captured and stored by an organism, and the stored energy is used to drive energy-requiring cellular processes. Photons are captured by antennae, where their energy is absorbed and used to extract electrons from water (water photolysis). These electrons then are channelled along electron carriers embedded in the thylakoid membranes and drive the exergonic pumping of protons from stroma to the thylakoid lumen (against the electrochemical gradient), generating a proton motive force across the thylakoid membrane. Protons then flow back to the stroma through the ATP synthases, thus driving ATP synthesis. During the oxidation steps, the electrons are used to reduce NAD(P)+ into NADPH. NADPH and ATP then drive carbon fixation and sugar anabolism in the Calvin cycle.
The overall reaction can be described by:

$$6\ CO_2 + 6\ H_2O \rightarrow C_6H_{12}O_6 + 6\ O_2$$

It is highly exergonic ($\Delta G^\circ = +2870$ kJ). Indeed, much of the free energy needed to drive this reaction comes from the oxidation of water. The subreactions that are dependent on sunlight to take place are referred to as “light reactions”, and they occur along the thylakoids. Water is the molecule of life and so it is pretty damn stable. Therefore, water oxidation is not an easy peasy reaction. The amazing feature of PSII is its ability to oxidise water and release molecular oxygen thanks to its oxygen-evolving complex. The water oxidation reaction has an oxidation potential of +0.82 V. As such, a strong oxidant is required to achieve this task. Fortunately, this is provided by the oxidised reaction centre chlorophyll P680+ in Photosystem II, which has an oxidation potential of +1.2 V and thus is able to do the job. The chemistry of water is inherently a 4 electron process, as four electrons must be extracted to evolve one molecules of oxygen. However, the photochemistry in the reaction centres takes place one electron at a time, so there is an apparent mismatch.

Fun fact: Structural studies revealed that Photosystem II in vivo is dimeric, in that two entire photosystems are closely associated forming a single large structural unit. Interestingly, there is little evidence that PSII is functionally dimeric, so each of the two reaction centres appears to function independently of the other.
This issue is solved by the storage of four oxidising equivalents by a single reaction centre working alone, which then oxidises water to oxygen in a concerted step. This is known as the S state hypothesis (developed by Kok and co-workers in the 1970s). Also we know that in PSII there is a manganese cluster in PSII. Later experiments confirmed that four Mn (which has 5 oxidation status and thus can take 4e- at a time directly from water and give them to P680+) directly take up 4 electrons from water in a concerted step.

The amazing process of light harvesting is carried out by pigment molecules. There is a remarkable number of pigments found in different photosynthetic organisms. Chlorophylls are the main ones, and they are squarish planar molecules, all containing

a magnesium ion coordinated by 4 nitrogens in a tetrapyrroles ring. Chlorophylls are classified chemically as chlorins. The prototypical chlorophyll is ChlA, which has an extensive delocalised pi electron system extending over most of the molecule. There is a tail (phytil tail) formed by condensation of four five-carbon isoprene units.
ChlA is found in all known eukaryotic photosynthetic organisms, together with ChlB. ChlB is identical to ChlA except at the C7 position, where it has a formyl group instead of a methyl. Usually the ratio of ChlA to ChlB is 3/1. Also ChlB has the absorption peak at lower wavelengths, thus have more energy (remember Plank). Interestingly, ChlB is found at higher concentrations on the outside of the reaction centres, and ChlA more inside. This is useful as ChlB can then channel the energy into the inside to ChlA and thus to the reaction centre.
There are other types of more exotic chlorophylls. Notably, diatoms and dinoflagellates contain ChlC, which does not have an isoprenoid tail (so it is classified as a porphyrin). Cyanobacteria have also ChlD and ChlF, which are the newest addition to the chlorophylls collection. These have huge biotech potential as they can absorb light in the Far-Red Light and Near-Infrared-Region of the electromagnetic spectrum, thus opening the possibility of expanding the photosynthetically active radiations (PAR).
Cyanobacteria are cool also because they have additional protein complexes called phicobilisomes, which coupe thousands of pigments to PSII. These have an allophycocyanin core with six radiating rods each containing 3 phycocyanin discs.

Engineering the light reactions
Photosynthesis is not very efficient. For example, only 25% of the total solar energy ends up in NADPH and ATP (during the light reactions. 13% ends up in sugars (dark reactions). And just 5% ends up in cell biomass after energy expenditure for growth maintenance and repair. Of course we need to remember that there is a trade-off between efficiency and robustness. The remarkable feature of living systems is that they are able to self-organise and self-repair, and this comes at the expenses of a lower efficiency.
However, one aspect of photosynthesis that can be optimised is its photon usage. Phototrophs usually absorb more energy than they require. This results in the saturation of productive photosynthesis at low light intensities. This might be an evolutionary strategy to deny light to competitors living below for example. The excess absorbed energy can then be non-photochemically quenched away as heat, as protection against overexcited states leasing to free radicals and oxidative damage. However, we don't care about evolutionary strategies if we want to produce as much energy as we can to increase the sustainability in the biosphere. In essence, there is scope for optimising the photon energy usage. Below listed are some strategies that have been demonstrated as effective solutions to increase photosynthetic efficiency during the light reactions.

1. **Accelerating recovery from photoprotection:**
   NPQ protects plants from excess light. However, the rate of NPQ relaxation is slower than that of NPQ induction. This asymmetry is exacerbated by prolonged or repeated exposure to excessive lights conditions. This slow rate of recovery of PSII antennae means that photosynthetic quantum yield of CO_2 fixation is transiently repressed upon a transition from high to low light intensity (so it is bad in field conditions of fluctuating light environment).
   Kromdijk et al., (2016) tried to accelerate NPQ recovery by manipulating the kinetics of the xanthophyll cycle. When leaves are exposed to high light, NPQ

is positively correlated with concentration of PsbS and stimulated by violaxanthin deepoxidation to zeaxanthin (catalysed by VDE). When light decreases, carbon fixing is limited as NPQ inhibits production of NADH and ATP. So rate of CO2 fixation is repressed until NPQ is complete, this can take minutes to hours and is correlated with the rate of zeaxanthin epoxidation, catalysed by ZEP. The group overexpressed PsbS, ZEP and VDE in tobacco, hypothesising that more ZEP would speed up NPQ relaxation, VDE would balance ZEP activity during NPQ induction and PsbS would adjust NPQ level to maintain WT amplitude.
Overexpression of these coding sequences from *Arabidopsis* into *Nicotiana tobacco*, resulted in an 15% increase in biomass under fluctuating light conditions.
Questions: they said they made single transformants with T-DNA integration. Is it in plastid genome or nucleus. If in nucleus, how can it work if cycle is in chloroplast?! Solved: the genes have chloroplast transit peptides…
Recently, Gomez et al. (2018) adopted a similar strategy but by focusing on Flvs protein, which relieve the excess of excitation energy on the photosynthetic electron transport chain by reducing oxygen directly to water. Introduction of cyanobacterial Flv1/Flv3 in tobacco chloroplasts resulted in transgenic plants that showed similar photosynthetic performance under steady-state illumination, but displayed faster recovery of various photosynthetic parameters, including electron transport and non-photochemical quenching during dark-light transitions. They also kept the electron transport chain in a more oxidized state and enhanced the proton motive force of dark-adapted leaves

There are often discrepancies in NPQ stidies maybe due to experimental errors. Ralph Bock and others recently presented evidence of an experimental artifact that may explain the discrepancies (van Oort et al., 2018): strong laser pulses lead to the formation of a novel electronic species in the major plant light-harvesting complex (LHCII). This species evolves from a high excited state of Chl *a* and is absent with weak laser pulses. It resembles an excitonically coupled heterodimer of Chl *a* and lutein (or other Xans at site L1) and acts as a de-excitation channel. Laser powers, and consequently amounts of artifact, vary strongly between NPQ studies, thereby explaining contradicting spectral signatures attributed to NPQ

2. **Optimise canopy shape:**
   Another strategy is to increase the area of leaves exposed to sunlight. Since productive saturates at low light intensities, the high canopy regions are going to be saturated in normal light and the lower levels are not actively photosynthesising. There are some mutants with a more spread-out leaf architecture and they could be interesting.
3. **Decrease antenna size:**
   Since phototrophs absorb more light than they need to, modelling suggested that decreasing antennae size could improve photosynthetic efficiency and enable to grow denser cultures (of both plants and cyanobacteria especially). Indeed, experiments reporting different levels of phicobilisomes truncation in Synechocystis showed that fewer antennae can lead to higher photosynthetic

efficiency. However, this comes at a cost, as microscopy imaging revealed that thylakoid membranes are significantly remodelled due to this modification and this may result in sicker cells, with lower growth rates.

4. **Extend the range of photosynthetically active radiation:**
   For most oxygenic phototrophs, usable light occurs between 400 and 700nm (PAR). Some algae and cyanobacteria are able to extend photosynthetic growth to 750 nm. This is significant as it allows them to live below other phototrophs, such as *Acaryochloris marina*. This is possible thanks to the use of red-shifted chlorophylls, ChlD and ChlF. They are identical to ChlA, except that they have a formyl group at the C3 and C2 position respectively. This formyl substitution shifts their absorption to 696 nm in ChlD and 705 nm in ChlF. ChlD is the major pigment in *Acaryochloris marin*a (which grows at the bottom layer in microbial mats). Yoneda et al. (2016) performed time-course RNA-transcriptomic experiments on *Acaryochloris* after nitrogen starvation. When nitrogen is not available, chlorophyll synthesis is inhibited and pigments are broken down (which causes chlorosis, and the loss of colours). As soon as the cultures are returned to a nitrogen-replete condition, the non-diazotrophic cyanobacteria resume chlorophyll biosynthesis. By doing RNA-seq you can identify the genes involved in ChlD biosynthesis. They found a lot of genes involved in general chlorophyll synthesis and nitrogen utilization but, most interestingly, a single oxygenase that followed the expression pattern of known chlorophyll biosynthesis genes. The name of the enzyme is impossible to remember but we can call it now chlD.
   To identify ChlF biosynthesis enzyme, Ho et al. (2016) used reverse genetics and heterologous expression of putative enzymes. Some cyanobacteria (like *Synechococcus* 7335 and *Chlorogloepsis fritschii)* are able to grow in far
   -red light and when they grow in FRL they undergo a photoacclimatation process (FaRLiP). FaRLiP involves the FRL-induced expression of a conserved cluster of 20 genes. They knocked-out those genes one by one in those bacteria and found out that psbA4 mutants could not produce ChlF. So they expressed it in the model cyanobacterium *Synechococcus 7002* and observed that it can synthesise ChlF. Thus, expression of psbA4 gene is sufficient to direct ChlF synthesis, showing that the gene encodes ChlF synthase, hereafter let's call it chlF! They propose that ChlF is a photo-oxidoreductase, that uses light to oxidise Chlide A to produce Chlide F, while reducing bound plastoquinone. Interestingly, this gene ("super-rogue") is a divergent paralog of the psbA genes encoding the D1 core subunit of PSII. Like PSII, ChlF uses light, ChlA, tyrosine Yz and plastoquinone but lacks the $Mn_4Ca_1O_5$ cluster. Phylogenetic analysis suggests that ChlF could be ancestral to PsbA of PSII, and PsbA could have arisen by gene duplication and divergence to bind the Mn cluster and catalyse water oxidation (instead of ChlA). Therefore, ChlF might be a long-sought, transitional intermediate in the evolution of oxygenic photosynthesis: a simple, homodimeric, type 2 reaction center that might have been evolved in anoxygenic ancestors of modern cyanobacteria.
   Importantly, introduction of ChlF biosynthesis into crop plans could expand their solar energy utilization. The group of Prof. Rutherford also showed that ChlF can drive oxygenic photosynthesis even if absorbs wavelengths with lower energy. A strategy could be to target ChlD to PSII and ChlF to PSI (as ChlD absorbs a shorter wavelengths and so higher energies). Or even replicate the organization of ChlA and ChlB around the antennae and target ChlD on the

outside borders of the reaction centers and ChlF inside, to channel the excitation towards the reaction centre.

5. **Artificial photosynthesis:**
   A different approach to use sunlight energy is to mimic photosynthesis outside of cells by converting sunlight into spatially separated electron/hole pairs within a photovoltaic cell and then capture the charges with catalysts that mediate water splitting. The four holes are captured by a catalyst at the anode to produce oxygen, and the four electrons are captured at the cathode to produce hydrogen. The net result would be the storage of solar energy in the chemical bonds of H_2 and O_2. Nocera et al., (2008) did this by making a catalyst that forms upon the oxidative polarization of an inert indium tin oxide electrode in phosphate-buffered water containing cobalt (II) ions. Amazingly, this catalyst not only forms in situ from earth-abundant materials but also operates in neutral water under ambient conditions.

The 'dark' reactions

The high-energy compounds (i.e. ATP and NADH) generated during light-induced electron transport along the thylakoid have intermediate stability. They are unsuitable for long-term storage of energy, such as building a plant. For this reason, these are used to drive the reduction of CO_2 into stable energy-storage compounds, as sugars and carbohydrates. The reactions that perform this anabolic task are referred to as "dark reactions", as they can take place also in the absence of light (but not exclusively). The metabolic pathway that incorporates inorganic carbon into sugars is known as the Calvin-Benson cycle (fun fact: they first used 2D paper chromatography and radioactive tracers to work out the pathway).
The pathway "begins" with ribulose 1,5 –bisphosphate and CO_2. For this reason, naively it has been proposed (e.g. by Donald Trump…) that global warming can indeed help in our mission of improving photosynthetic efficiency. If by burning a shit ton of fossil fuels we release more carbon dioxide, since CO2 it's a reactant, by the law of mass action this would result in a "fertilization" effect to increase biomass. Actually, the law of mass action is true and this would increase the rate of photosynthesis. However, the problem is that by increasing carbon dioxide levels, the temperature also increases and this results in drier land and increased photorespiration rate. So the fertilization effect is counterbalanced by an increase in photorespiration and thus a decrease in photosynthetic efficiency… This is why we need to engineer more efficient and resilient crops, especially bearing in mind the high temperature and high carbon dioxide levels predicted for the future.

One molecule of glucose is produced every 6 turns of the Calvin cycle (6 CO_2). Overall the carbon fixation reaction can be described by:
$6CO_2 + 18ATP + 12NADPH + 12H_2O \rightarrow C_6H_{12}O_6 + 18ADP + 18P_i + 12NADP^+ + 6\ H^+$
The overall cycle can be broken down in three phases: carboxylation, reduction and regeneration. The carboxylation step generates PGA. The reduction phase uses the NADPH and some ATP produced by light reactions to reduce PGA to triose phosphate. Most of the triose phosphate is used to regenerate RuBP in a complex set of reactions, while some of it is drawn off for starch synthesis in the chloroplast and sucrose synthesis in cytoplasm. Most of the reactions in the pathway resemble that of

gluconeogenesis and the oxidative pentose phosphate cycle. Also the enzymes catalysing these reactions are homologous but often differ in the type of regulation. Regulation often occurs via thioredoxin-mediated redox control at Cys residues. The unique reactions are the carboxylation and the final step in the regeneration of RuBP. Rubisco (ribulose-bisphosphate carboxylase/oxygenase) is the enzyme that carries the carboxylation step. An interesting random fact is that the reaction catalysed by Rubisco in cyanobacteria is the only anabolic reaction in the biosphere that takes place in a proteinaceous compartment (carboxysomes). Rubisco from higher plants consists of eight copies each of large (L) and small (S) subunits, giving a L_8S_8 quaternary structure. The L subunit contains the catalytic site, shared between two L subunits (so minimal functional complex is a L_2 dimer). The function of the small subunit is still not well understood, and as usual for non-understood processes, we say it might be involved in stabilization or regulation. In most of the photosynthetic eukaryotes, the large subunit is encoded in the chloroplast genome, whereas the small one in the nucleus. Chaperonins are essential for Rubisco assembly.
The carboxylation is very exergonic ($\Delta G° = +35$ kJmol^{-1}), so it is essentially an irreversible reaction. The mechanism is thought to involve spontaneous formation of an enediol of RuBP (facilitated by coordination of the magnesium ion). The enediol formation is followed by carboxylation, to give an unstable intermediate that breaks down into two molecules of PGA (one of which contains the carbon atom that was derived from atmospheric CO_2).

What is remarkable about the Calvin cycle is that it uncouples fixation from reduction. It is not that hard to fix carbon, what is hard is fixing it and not losing any of it along the pathway. Also the cycle is autocatalytic, it produces its own substrates. Is this an example of emergent self-organising network? (perhaps the most important one, which gave birth to our life on earth).

However, in addition to carboxylation, Rubisco can also catalyse the oxygenation of RuBP, giving one PGA and one molecule of 2-phosphoglycolate instead of 2 PGA (called photorespiration). This makes Rubisco not very efficient and so a lot of bad things have been said to the poor Rubisco (“Rubisco is slow and confused”). Also, it has a very slow catalytic rate (of circa 3 reactions per second) and so it has to be highly expressed in leaves (causing a lot of burden and allocation of resources to that). The oxygenation reaction in particular is the single greatest step in plant metabolism, as it happens approximately a third of the time (and the frequency increases at increasing temperatures). CO_2 and O_2 are competitive substrates, for this reason the ration of carboxylation to oxygenation is determined by the catalytic rates and affinities of Rubisco for the two substrates. We can calculate the specificity factor τ (kinetic preference for CO_2 over O_2 = $(K_{cat}^{CO2}/K_m^{CO2})$ / (K_{cat}^{O2}/K_m^{O2}). Although this number indicates a high preferment for CO_2 over O_2, higher plants Rubiscos have significant oxygenation activity because of the higher concentration of oxygen in the atmosphere rather than carbon dioxide.
The problem of oxygenation is that it produces the toxic intermediate 2PG, which in turn has to be recycled into phosphoglycerate by photorespiration. This costs ATP, CO_2, NADPH and NH_3 (especially by glycine decarboxylase). The recovery involves also three organelles to split the heavy burden of this recycling job, the chloroplast, mitochondrion and peroxisome. The question though is why photorespiration in the first place? Think about the evolutionary reasons of photorespiration. Nogales et al.

(2012) performed a systems biology analysis and modelling results actually showed that photorespiration is essential for high light acclimation and optimal photosynthesis. These results suggest that Rubisco's oxygenation is a result of a compromise between chemical and metabolic imperatives.

Engineering the 'dark' reactions

Rubisco is the single greatest loss step in plant metabolism. For this reason, one of the "Holy Grails" of photosynthesis research is to engineer Rubisco or the dark reactions to improve CO_2 fixation, as a strategy to increase crop yield.

1. **Heat-resistant Rubisco Activase**
   Before Rubisco can become catalytically active, a lysine residue must be modified by carbamylation, by Rubisco Activase. One CO_2 molecule reacts with the ε-amino group of a specific lysine in the L subunit. A Mg^{2+} ion coordinates to the carbamylated residue to activate the enzyme. The carbamate is thought to act as a general base, facilitating the formation of the enediol intermediate. If a sugar binds to the non carbamylated Rubisco, the enzyme becomes inhibited and no reaction takes place. Rubisco activase activates and unstucks it. Interestingly rubisco activase is also regulated by thioredoxin. However, the activase is very heat sensitive. For this reason, it has been proposed to engineer a more heat resistant activase. Kurek et al. (2007) improved the thermostability of rubisco activase by gene shuffling. Experimental results demonstrated that indeed rubisco activase is a rate limiting factor under moderately high temperatures and evolving a more heat-resistant rubisco activase can improve resistance to heat stress. Something to bear in mind as temperatures are currently higher than historical averages and are predicted to increase even more in the near future.
2. **Improving the regeneration step**
   Even though most of the time Rubisco is the rate limiting step of the Calvin cycle, sometimes the regeneration steps become limiting. It has been shown that playing with the regeneration kinetics by upregulating the enzyme SBPase resulted in increased growth in tobacco (Rosenthal et al., 2011). Another strategy could be to reallocate the nitrogen budget (Zhu et al., 2007).
3. **Improving rubisco**
   Another strategy is engineering rubisco to increase its Kcat or specificity for CO_2. However, apparently there is a trade-off between efficiency and selectivity and it is hard to modulate the two parameters independently. Lin et al. (2014) replaced the chloroplast rubisco large subunit in tobacco with an operon encoding small and large subunits from *S. elongatus* (as cyanobacterial rubisco have a higher Kcat). It showed increased efficiency, however there was poor assembly and it only worked at very high CO_2 concentrations (9000 ppm vs the natural 4000 ppm). More recently, the same group used a similar approach to express the two rubisco subunits from the red alga *Griffithsia monilis* in chloroplast tobacco (Lin et al., 2018). Although the red algal Rubisco genes are being transcribed in tobacco chloroplasts, the transgenic plants lack functional Rubisco and can only grow in a medium containing sucrose. Their results suggest that co-expression of compatible

chaperones will be necessary for successful assembly of red algal Rubisco in plants.
Occhialini et al. (2016) then demonstrated the you can avoid insertion of GroEL and GroES interactions with Rbxx, to introduce a simpler version. In essence, this strategy is to survey rubisco biodiversity and improving step by step the parameters.

4. Implementing carbon-concentrating mechanisms

Tropical plants and cyanobacteria came up with clever strategies to overcome the poor selectivity of Rubisco by evolving carbon-concentration mechanisms, so that rubisco is only surrounded by CO2 and the oxygenation reaction does not take place by simple law of mass action.
Certain plants that live in hot climates (including maize and sugarcane) are C4 plants and they can temporarily fix CO_2 in certain cell types and the resultant organic carbon compound is transported to another cell type, where is decarboxylated and so there is a lot of CO_2 liberated and concentrated around rubisco, avoiding photorespiration. This mechanism is not a replacement to the Calvin Cycle but an addition to it, but of course it comes with higher cost in terms of ATP (which however is less than the energetic price paid due to photorespiration). C4 plants are so named as carbon is first fixed into a 4–carbon compound (oxaloacetate) instead of a 3-C (3PG).C4 plants are anatomically different, as they have a characteristic circle of cells surrounding the vascular bundle called bundle sheath cells. Around bundle sheath cells (towards the outside) there are mesophyll cells. These cells are connected by very specific pores called plasmodesmata (which allow exchange of metabolites). The primary fixation od CO_2 takes place in mesophyll cells, where PEP carboxylase catalyses the carboxylation of HCO_3^- with PEP to form oxaloacetate. The substrate of PEP carboxylase is bicarbonate (formed by the action of carbonic anhydrase) and not CO_2, so O_2 is a poor competitive substrate and the oxygenation reaction does not take place. Oxaloacetate is the reduced into malate and transported to the bundle sheath cells, where it is decarboxylated and converted to pyruvate. Pyruvate is transported back to the mesophyll where it is phosphorylated using ATP. The CO_2 liberated in the bundle sheath cells then enters the classic Calvin cycle. An additional advantage to less photorespiration, is that since C4 plants are better at scavenging CO_2, they don't need to keep the stomata open as much as C3 plants, and so they also suffer less from water loss.
Interestingly, it seems that C4 metabolism evolved independently multiple times in multiple plants, and also some C3 are "preadapted" to C4 metabolism, so it should be easy to engineer in C3 plants, right?! The problem is that there are large anatomical differences between the wo types of plant, which are hard to genetically engineer.
A similar but different strategy has been adopted by cyanobacteria and some algae. This consist of a series of pumping mechanisms, clever positioning of carbonic anhydrase and dense packing of rubisco inside proteinaceous compartments called carboxysomes. The carboxysome shell consists of three structurally characterized protein types, each named after the oligomer they form: BMC-H (hexamer), BMC-P (pentamer), and BMC-T (trimer). These three protein types form cyclic homooligomers with pores at the center of symmetry that enable metabolite transport across the shell. Carboxysome shells contain

multiple BMC-H paralogs, each with distinctly conserved residues surrounding the pore, which are assumed to be associated with specific metabolites

There are two types of carboxysomes, distinguished by the form of RuBisCO they encapsulate: Form IA in □-carboxysomes and Form IB in β-carboxysomes. Accordingly, cyanobacterial species encoding each carboxysome type were termed □- and β-cyanobacteria. □-cyanobacterial strains occupy the open ocean and their carboxysomes are encoded in a single genomic locus. In contrast, the β-cyanobacteria inhabit a range of dynamic habitats. Whereas most β-carboxysome genes, including the genes encoding the BMC-H proteins CcmK1 and CcmK2, are situated in the main carboxysome locus and are constitutively expressed across a wide range of conditions, all β-cyanobacterial genomes also contain varying numbers of additional CcmK paralogs (CcmK3-CcmK6) in satellite loci. It has been hypothesized that the different CcmK paralogs have differing metabolite selectivities.

The Kerfeld group recently investigated the function of the satellite locus-encoded proteins CcmK3 and CcmK4. Growth analysis of deletion mutants shows that CcmK3 and CcmK4 are not redundant and that CcmK3 is a component of the β-carboxysome that is expendable under ideal growth conditions. CcmK3 does not form homohexamers when expressed recombinantly, but CcmK3 and CcmK4 form a heterohexameric complex that can further form dodecamers under certain conditions (Summer et al., 2019).

Beta-carboxsomes are one of the few BMCs that have some of its genes in satellite genome loci. This might be important because beta-carboxisome is the only known BMC that is constitutively expressed all the time. Having some of its genes in distal loci may enable differential expression. They suggest that CcmK4 can form both homohexamers and heterohexamers with CcmK3, and the two types of hexamers have distinct permeability properties.Thus heterohexamer formation could provide a means of fine-tuning β-carboxysome shell permeability.

CO_2 or HCO_3^- is pumped into the cell by ATP-driven pumps, a Na^+/ HCO_3^- symporter and a CO_2 transporter. The advantage of HCO_3^- is that is very membrane-impermeable so it does not leak out. Carbonic anhydrase in the carboxysome which converts bicarbonate into CO_2, which is immediately fixed by rubisco before it can leak out. Interestingly there are not carbonic anhydrase in the cytoplasm of cyanobacteria and if you introduce them, cells are not able to concentrate CO_2, this is because otherwise it bicarbonate would be reconverted back into CO_2 and leak out of the cell.
Long et al., (2018) recently reported the addition of 4 genes from the *Cyanobium* cyanobacterium in tobacco chloroplasts. Interestingly, this is the most minimal carboxysome construct ever reported (only Rubisco small and large subunit and 2 carboxysomal self-assembling proteins). They introduced them into a multigene operon with intercistronic elements and all the standard transplantomic package… This resulted in increased photosynthetic efficiency. Algae do a similar job but by using chloroplast complexes called pyrenoid bodies, which are essentially crystalline rubisco with carbonic anhydrase.

5. Photorespiratory bypasses

Biochemical constraints as well as abiotic factors are crucial to bear in mind when trying to engineer Rubisco. Because of this complexity, another strategy is to accept the flaws of Rubisco and try to recover the damages (or even exploit) of photorespiration, by engineering Photorespiratory bypasses.
Kebeish et al. (2007) reported the transfer *E. coli* glycolate pathway into chloroplasts (5 enzymes). The aim was to bypass photorespiration by assimilating the glyoxylate produced by photorespiration using a bacteria pathway. However, the problem with this is that yes it circumvents the loss of nitrogen but the glyoxylate reaction still results in loss of CO_2. In addition, the transformants expressing only the first enzyme of the pathway showed similar results, which make the experiment a bit controversial and debatable.
The Kerfeld group at Berkeley proposed as an alternative strategy to introduce the right-hand side of the 3-hydroxyproprionate cycle. This is a C-fixing bicycle discovered in *Chloroflexus auranticus*. This cyanobacterium, fixes bicarbonate with a biotin-dependent acetyl-CoA carboxylase and propionyl-CoA carboxylase. The primary carbon fixation product of this cycle is glyoxylate (the same one produced by photorespiration). In the 3HP cycle then, glyoxylate is fed into the second cycle of the bicycle, where a new bicarbonate is fixed and pyruvate is generated as final product. Therefore, if you introduce the 3-HP cycle genes of the right end side, you can use glyoxylate waste product of photorespiration, save it and also fix one more C at the same time. They introduced the 6 genes in *S. elongatus* and this resulted in no NH3 loss and net in in carbon fixation compared to C2 cycle. For the future, it is important however to make cyanobacteria more resistant to 3HP, as it is a bit toxic. Also, the 3HP pathway interacts with fatty-acid biosynthesis (malonyl-CoA metabolism), which could distort this essential chloroplastic process.

South *et al.* (2019) very recently reported a newly designed photorespiratory bypass pathway in genetically modified tobacco plants.
In this pathway (synthetic glycolate), only two transgenes have to be introduced into the plant chloroplast: a glycolate dehydrogenase that converts glycolate into glyoxylate derived from the green alga *Chlamydomonas reinhardtii* was redirected to tobacco chloroplasts, and a malate synthase was expressed to convert glyoxylate to malate and eventually to acetyl-CoA via the native chloroplast-resident nicotinamide adenine dinucleotide phosphate (NADP)–malic enzyme. Using the green algal glycolate dehydrogenase instead of plant glycolate oxidase prevents production of hydrogen peroxide, and hence additional expression of catalase is unnecessary.
Besides introducing a synthetic bypass, South *et al.* also reduced the expression of PLASTIDIAL GLYCOLATE/GLYCERATE TRANSPORTER 1 (PLGG1). This modification was suggested previously to increase the potential of synthetic bypasses, because it restricts the export of glycolate from chloroplasts and hence promotes its consumption by the synthetic bypass. A larger portion of glycolate is decarboxylated within the chloroplast by the synthetically engineered bypass, leading to enhanced CO_2 fixation activity of RuBisCO. This comes with an impressive yield gain of more than 40%.

6. **Implement other autotrophic C-fixing pathways**
   Of course in nature there are other carbon fixing pathways in addition to the Calvin. For example, green sulfur bacteria run the TCA cycle backwards. They used reduced ferredoxin produced by the reaction centre to reduce carbon dioxide, producing pyruvate and acetyl-CoA.
   Other pathways include the reductive Acetyl-CoA pathway, the hydroxyproprionate-butyrate pathways etc.. However, with all these pathways there are always problems with oxygen sensitivity.
   The most promising alternative is perhaps that adopted by filamentous anoxygenic phototrophs (like *C. auranticus*), that use the 3HP cycle to fix carbon. Flux Balance Analysis suggests that is feasible to introduce it in a cyanobacterium (only 6 genes required). The 3HP is also the modelled pathway that gives the best growth rate (5% better than WT) at 5% PHOTOR, among the feasible pathways.
   However, if we want to introduce it in plants there is a problem with an enzyme, methyl-malonyl-CoA mutase, which requires cobalamin as a cofactor. Plants do not produce cobalamin and so it would not work there. A solution could be to do directed evolution on this enzyme and evolve it to use a different cofactor than cobalamin.
7. **Synthetic C-fixing pathway**
   A radical alternative is to forget Rubisco altogether and natural pathways and engineer carbon fixing pathways de novo, from the bottom up, using computational models**.** There is nothing special about carboxylation, the problem is with the self-organization and sustainment of natural pathways.
   In silico studies revealed several candidates. The most promising are the MOG and CETCH pathways. The MOG (Malonyl-CoA-Oxaloacetate-Glyoxylate) cycle uses PEP carboxylase which is the best carboxylating enzymes in terms of Kcat and Kcat/Km.
   Recently, Schwander et al. (2016) constructed and introduced the synthetic CETH pathway in vitro. This uses 17 enzymes including 3 engineered from 9 different organisms. The pathway uses reductive carboxylation to add CO_2 onto metabolites and was measured to be up to five times more efficient than the in vivo rates of the Calvin cycle.
   However, both the MOG cycles and the CETCH cycle produce glyoxylate, the assimilation of which into central metabolism via the bacterial glycerate pathway is rather inefficient as it involves a decarboxylation step. Maybe it would be good to add one of the synthetic pathways in addition to the 3HP cycle?
   Moreover, these cycles are very long and highly complicated. While implementing these pathways in microbes might be possible, their integration into the plant metabolic network seems highly unlikely.

Plant Biotechnology - Syngenta

Agriculture and Sustainability

There is a lot of concern regarding food security in the coming years. In particular, the global population is expected to reach 9 billion by 2050. This corresponds to an increase in 50% of calories that need to be provided if we want to sustain this growth. As a result, the agricultural market needs to face the challenge of increasing crop yields by 50% by 2050. This mission is complicated by the fact that we have limited availability of arable land, and as a result to the increase in population, arable land is going to decrease in the future. Water is one of the main substrate of photosynthesis and as such is essential to gro crops. Because this is finite and limited resource, we cannot afford to increase crop yields by linearly increasing water usage. Another challenge is the potential threat caused by ecological issues, as crops are major components of the ecology of the world; and any unexpected imbalances in natural ecosystems may lead to catastrophic events. In essence, we must increase productivity but in a sustainable way. A sustainable agriculture is one that preserves biodiversity, avoids deforestation, protects health and safety of workers and end-consumers, protects the economic viability of farming, ensures soil conservation, maintains water quality and suits local conditions and capacity. Precision agriculture and biotechnology are exciting fields, with the potential to unlock plant potential through innovation.

The green revolution and the advent of biotechnology in last decades have already enabled a remarkable increase in crop productivity. For example, now we can grow 1 ton of corn using 37% less land. In particular, four major technologies helped to reach this increased productivity:

1. Mechanization (e.g. irrigation)
2. Synthetic fertilisers (e.g. NPK)
3. Crop protection chemicals
4. Better seeds (via traditional breeding, marker-assisted breeding and GMOs.)

Two main parameters have been significantly optimised by the wave of innovation brought by the Green Revolution: the probability of radiation interception and the biomass partition efficiency. For example, new traits that confer less branches in corn enable the leaves to be irradiated by more sunlight, thus resulting in increased photosynthetic quantum yield. Similarly, plants with less shoots and stems have been selected/engineer so that fixed carbon ends up more in edible biomass (e.g. fruits and leaves) rather than in waste products (e.g. stems). Oher good innovations, even though not strictly novel/exciting (my own personal view) are crop rotations, no-tilling practices to maintain soil quality. Or even preserving the wilderness/biodiversity of beneficial insects, pollinators and birds via integrated pest management strategies.
However, this agricultural revolution was bypasses in certain regions, such as in Sub-Saharan Africa. As a result, there is a lot of untapped potential for AgTech.

Several exciting AgTech developments are being developed recently. These include the use of positioning technology (e.g. GPS) to plant seeds with 2 cm accuracy (and then the sprayer knows where to spray precisely), sensor technology to monitor plants remotely, increased computing power/robotics/machine learning to detect diseased plants with drones, better delivery systems, big genomics data, advances in molecular biology, and alternative protection systems (e.g. behaviour modifying chemicals).
Start-up Idea: all of this is exciting, however there is still a big missing link : we still do not have a system for precision application of weed control (we actually have to take away weeds by hand or spray by hand…).

Regulation, Food Safety and Security
If we want to make/sell agrochemicals we must demonstrate that these are safe for consumption. There is a lot of research/regulation to do this, and it is quite similar to the pharma industry.
However, there is a gap between the expert's assessment of risk and the perception of risk by the general public. For example, one of the main assessment valued by expert's regulation is the probability of microbial contamination or nutritional imbalance. The public on the other side is scared that products contain traces of food additives or residual pesticides.
We need to remember that a given risk can be quantified by its intrinsic hazard (e.g. forms and modes of toxicity of a given pesticide) summed to the exposure of that particular hazard (e.g. application rate, movement and dispersion etc…).
It is the dose that makes the poison. To measure hazard, we can plot dose-response curves. The highest concentration of a given chemical at which there is no observed adverse effect (that is the half-height of the sigmoidal dose-response curve) is called the NOAEL. The highest concentration at which there is no observed effect is called the NOEL. The acceptable daily intake (ADI) of a given molecule can be calculated by dividing the NOEL by 1000.

$$ADI = NOEL/1000$$

In USA and Canada there is a risk-based assessment, where hazard and exposure are considered. For this reason, the ADI is calculated by dividing the NOEL instead of 1000 by a safety factor proportional to the hazard posed by a given chemical.
In Europe, concentrations are only by exposure cut-off values. For example, pesticides cannot be found at concentrations higher than 0.1 ug/l in drinking and ground waters, independently of the hazard posed by that specific molecule.

Surprisingly, modern insecticides have similar toxicity to caffeine, and 99.9 % of pesticides in our diets are phytochemicals (natural plant products). For example, p-hydrazinobenzoate, D-limonene and sesamol are all potent carcinogens found in relatively high concentrations in mushrooms, oranges and sesame seeds, respectively.
So clearly there is a fundamental question here. What does it mean to be natural? Let's get away from the concept of "natural", "organic", "bio", everything is chemicals and let's embrace a quantifiable definition.
Still now we classify crop protection products as natural or synthetic. These chemicals then get broken down by sunlight and other biochemicophysical ways and their residues may end up in the final edible product. So the risk

depends on the concentration of the residue times the intrinsic hazard of the chemical species. If there is no residue, there is no exposure. It is as simple as that.

Now, suppose that I have a biochemical company that produces pesticides. In order to put my product in the marker I need to get a pesticide registration. To achieve this document, I need to prove that my chemical does not pose any unacceptable risk to humans (operators, consumer, etc…) and the environment (aquatic systems, birds, bees, soil, etc…). To demonstrate this, I need to carry investigative studies, ranging from toxicology studies (short + long term effects, mutagenicity, developmental neurotoxicity, oral LD50 etc…), residue studies, ecotoxicologicity. An important one is the dietary exposure. In EU, there is a maximum residue level that needs to be quantified in order to check is products are not misuses. This is calculated as the mg of residue over the Kg of crops. Crops cannot be found with values higher hat this (it changes from pesticide o pesticide).

Food safety is very important; this can cause a lot of damage. For example, *Aspergillus flavus* is a Fusarium fungus that produces the mycotoxin aflatoxin, which causes liver cancer and death if taken in high risk. This fungus infects cereals and its contamination costs 200$ Million every year. For this reason, its spread is carefully monitored. However, there are no fungicides available for this type of contamination. There is one solution: Afla-Guard is a non-toxin producing strain of *A. flavus*. If this is applied to the crop, it can outcompete the toxin-producing fungus and prevent contamination. A very elegant and bio-based solution indeed.

Ever-more powerful genetic technologies, such as genome-editing endonucleases and marker-assisted breeding, continue to facilitate the development of genetically modified (GM) crops engineered with complex traits, such as, nutritional quality, climatic resilience and stacked disease-tolerance mechanisms. But in many developing countries, the uptake of these GM products is being jeopardized by the sluggish pace and inadequacy of regulatory oversight. This is a serious concern because developing countries stand to benefit most from the adoption of new varieties of staple GM crops, such as vitamin-enhanced rice and bananas or disease-resistant maize and cassava. Despite the availability of the formal risk analysis framework—which provides all the critical components of risk assessment, risk management and risk communication important for structured regulatory decision making on such products, policymakers do not always understand the underlying factors behind a risk analysis well enough to facilitate implementation of robust and realistic biosafety practices
At the heart of the problem is a lack of agreement as to whether and how both scientific and non-scientific evidence can and should be integrated into regulatory decision-making for GM crops. The risk analysis framework embodied in the International Plant Protection Convention, the Codex Alimentarius (Codex) and the World Trade Organization (WTO; Geneva) is based solidly on science. In contrast, the precautionary principle embedded

in the UN Cartagena Protocol on Biosafety balances scientific evidence with economic, social and environmental norms.
The main problem is that decisions in Europe are often made on political grounds, rather than on a scientific basis.
One possible way forward would be to include assessment of socioeconomic considerations only when there is clear evidence of the socioeconomic changes that would result from the introduction of the GMO.

Crop Biotechnology

One of the main hype about Biotech for agriculture is because it allows to insert specific desired traits in a very precise way into crop plants. From insect-resistance, to increased heat tolerance, to better nutritional values, traits can be discovered and inserted into commercial cultivars.
It is possible to differentiate between input and output traits. Input traits are those that favour the agribusiness and growers. These include modifications that lower production costs, increase yields and promote no till practices, such as insect/herbicide resistance traits, stress tolerance etc…
Output traits add new sources of values for retailers or end consumers. These include traits with modified purpose (e.g. biofuels, feed quality, bio-remediation), traits that transform plants in biofactories (chemical feedstocks, enzyme production, biopharma), and traits for better food (improved nutrition, taste and health).

Domestication of plants for human purposes is nothing new. We have always modified nature for our own benefits. However, whereas traditional breeding strategies were unconscious and phenotypic, biotechnology allows to select for new traits consciously and genotypically.

In order to develop a new GM crop. First we need to find a source of useful gene. Once a useful gene has been identified, it has to be isolated (usually though PCR amplification) and put into a vector. The vector is then used to transform a plant tissue. There are several transformation techniques, including biolistic bombardment, protoplastic and Agrobacterium-mediated (developed by Mary-Dell Chilton, founder of Syngenta Biotech) transformation. A hot field is now also transplantomic, where genes are inserted into chloroplasts (which have way more DNA copies and thus way more product is expressed (and lots of other exciting advantages).
After transformation, you need to select for the desired trait. Most common is insertion of electable marker manA from E. coli, which allows the transformed tissue to convert Mannose-6-Phosphate into Fructose-6-Phosphate and thus grown on an alternative carbon source.
After transformation and selection, the tissue needs to be regenerated.

One of the first trait to be inserted into GM crops was the ability to grow in environment with herbicides, so that Monsanto could sell a shit ton of Round-Up herbicide and crops would still survive. Together with Ringspot-resistant Sunrise Papaya, Round-Up ready (Monsanto) and Liberty-link (BCS) crop lines were the first to be commercialised.

Liberty-link crops are resistant to the herbicide glufusinate (phosphoinothricin). These plants have been transformed with a bar gene from *S. hygroscopicus*, which is a phosphoinothricin N-acetyl transferase (PAT) enzymes, catalysing the acetylation and hence inactivation of the glufusinate herbicide.
On the other hand, Monsanto's round-up ready crops are resistant to glyphosates, which inhibit aromatic amino acid synthesis via inhibition of a key enzyme (EPSPS) in the Shikimate pathway. The GM crops have been transformed with a modified version of the EPSPS enzyme, which has a modification in the active site which renders it insensitive to glyphosates.

Another common GM trait is herbicide resistance. The most famous example of insect-resistant GM crop is Bt corn. *Bacillus thuringiensis* is a prokaryote that synthetises a Cry δ-endotoxin, Bt toxin. This has been used for year (and it is still used) as a "natural" insecticide as it ruptures the cell lining the gut of Lepidoptera, common corn pests. Bt corn expresses this toxin, which kills the insects when they eat the GM crop. Importantly, it is only activated in the very high pH of Lepidoptera and this it is harmless in humans. Interestingly, this Bt toxin, if sprayed, is considered an organic insecticide, but when inserted in the genome is the end of the world… What is morality? What is "natural"?...
Now they also modified the enzyme to kill Coleoptera (beetles).

Other cool traits include engineering drought tolerance (by engineering ABA receptors) and many more (see below). Very hot is also the use of RNAi as alternative insecticide system. If you spray onto leaves some dsRNA that when cleaved by DICER binds and inactivates insect RNA coding for some vital function, then you have your "natural" insecticide. Amazingly, this has been shown to work in field conditions with just topical applications of dsRNA. Monsanto recently claimed that they can reverse Round-Up resistance in Lepidoptera using RNAi. But words are just words, we want papers! The only problem is that it is quite expensive to synthesise dsRNA as insecticide!
Why not expressing it in the chloroplast directly?

Extra examples of cool GM traits:

- Edible cottonseeds using RNAi to decrease gossypol, which is toxic to human and in high concentration in cottonseeds. (Waltz, 2018)

- De novo domestication with CRISPR multiplexing of Tomato to confer day-length sensitivity, compact shoot and flower architecture (for high-density growth and mechanical harvesting), synchronised fruit ripening, enlarged fruit size, and extra ascorbic acid biosynthesis. (Li et al., 2018).

- Combining feedback control theory with synthetic biology for engineering plants with enhanced resilience to environmental stress. Designed and tested perturbation mitigation strategies based on the use of genetic feedback control and showed how a synthetic feedback controller can be designed to attenuate the effect of external perturbations due to fungal infection in Arabidopsis. (Foo et al., 2018).

- The branched-chain amino acid (BCAA) biosynthetic pathway performs 'house-keeping' functions in plants and fungi but is non-existent in animals. Yan et al. (2018) demonstrated that developing compounds which interrupt the BCAA biosynthetic pathway might be an efficient approach to eliminate plant growth without side effects in animals.

  And since we are talking about hyped stuff, why not mentioning CRISPR?
  Of course, all the overheard qualities of CRISPR apply also to GM crops: ease of use, specificity, cheap etc… Importantly there are two ways in which you can use CRISPR in plants: DNA-based (inject plasmids with the three components – Cas9, gRNA, HR-template) or Protein/RNA-based (inject Cas9 protein + sgRNA).

  Native Traits and Plant Breeding
  Breeding is the main strategy to increase plant biodiversity. Biodiversity drives crop potential.
  Gene mining provides allele diversity. New alleles with interesting traits are then delivered into chassis cultivars. Hybrid lines are crossed until candidate cultivars are developed. Candidate cultivars can then evolve into commercial cultivars, depending on market parameters (so not discussed here, as this is not bullshit).
  New traits/alleles are identified by reverse or forward genetics. The advent of Next-Generation-Sequencing facilitates this process. Then genotypic differences are mapped to phenotypic differences. However, nature does not provide us with all the traits we need. So mutagenesis is very important. Random mutagenesis by ethylmethosulfonate (EMS) was the standard random mutagenesis workflow. However, EMS is nasty and carcinogenic and the process is not very efficient. Now directed evolution is the way to go.
  Cloning is needed to insert the gene of interest. Cloning can be homology-based, by complementary genetics, map-based or directly genome editing.
  When you insert the transgene, biochemical knowledge is required and systems biology approaches are beneficial to understand the effect at a holistic level.
  Since a long time ago (actually Darwin was one of the first to observe this phenomenon) F1 hybrids have been known to be fitter than the parental lines. This phenomenon (known as 'hybrid vigor') is probably due to increased heterozygosity of alleles, however the mechanistic details are still not fully elucidated. F1 progenies are phenotypically identical but genotypically dissimilar. So usually farmers buy F1 hybrids to improve crops but then they cannot grow the F2 progeny as it would result in different phenotypes. So usually they have to buy other F1 seeds. When a new F1 hybrid is developed it has to be fixed, so that the genotype can be always the same and you can sell that plant. The classic way to fix the F1 genotype is though several rounds of backcrossing via self-pollination. This process is long (up to 6 years9 and results in ca. 96% homozygosity.
  Double haploid technology (DH) is a more effective and quicker solution to achieve F1 homozygosity. To achieve this, a diploid female is crossed with a male inducer (if we are doing wheat DH then it's maize male). Because of the high genetic homology, maize pollen can pollinate wheat receptive anthers, however the maize chromosomes are then eliminated by the wheat. This

results in a haploid embryo with the mother parental genotype. Then colchicine is applied to induce chromosome doubling, resulting in a doubled haploid line, homozygous at all loci.
DH technology has several advantages, including:
1. Shortened breeding process (1 year)
2. Increased DH uniformity (100% homozygosity)

There are three main breeding strategies:
1. Conventional: driven by phenotypic selection
   Not very rational/high-throughput
2. Marker-Assisted: driven by molecular selection
   It requires a linkage map- Marker-based trait screening. There is a problem with linkage-drag of donor chromosomes, the markers can reduce the size of this dragged DNA.
3. Genome Editing: driven by genetic selection
   What can we say about this, if not rational and genotypic?

Discovery of Agrochemicals

Crop protection is a major factor driving crop productivity. Agrochemicals include insecticides, fungicides, herbicides and all contribute to higher yields, better-quality products, reliability and ease of harvest.
Agrochemicals are nothing new. Around the 2500BC, Sumerians reported the use of sulphur to control crop pests. Then several inorganic compounds have been experimented with some (poor) results. In the 1940 the first synthetic insecticide was developed: DTT. Since then there has been an explosion in better organic synthetic compounds, including paraquat (herbicide), glyphosate (herbicide), mesotrione (herbicide), solatenol (fungicide, 2010).

Oe of the main target is weed. They significantly compete with crops for sunlight, nutrient and water, negatively impacting yields. In the UK, blackgrass is a major pest for cereals. In the US, Amaranthus is a superweed the negatively affects corn growth.
Herbicides have provided many socio-economic benefits. Before, workers needed to pick weeds by hand (some still do). Use of herbicide favoured moving from land to cities, urbanization, and industrialization.
Herbicides can be selective or non-selective (glyphosate). The more selective the better.
Fungicides are very important as fungi are major pests (the great Irish famine caused 1 millions of deaths and was caused by late potato blight contamination). Insect pests can be chewing (e.g. caterpillars), which decrease the photosynthetic area) or sucking (e.g. aphids), which can transmit several nasty viral infections. Nematodes are also very bad because they cause severe damages to root systems.

The process of finding a new agrochemical is very much similar to the pharmaceutical industry. It starts with the generation of a lead compound, followed by optimization, development, registration and finally may result in a commercial pesticide after a 10 years long process costing hundreds of millions of dollars.

The main difference is that here the screening occurs in vivo, which facilitates the all thing.
The mode of action of the chemical must deliver rapid and self-sustaining damage in a single application (wave-spike curve). That is there must be an irreversible lethal effect within a window of persistence.
Sources of new leads are diverse, from serendipity to structure-based design.

Chemicals need to have certain chimico-physical properties in order to ensure appropriate transfer process, absorption, translocation, root uptake, shoot/leaf uptake and biological effect. They have to be resistant to light, oxidative stress, rains etc…
The optimization step is performed using the iterative Design-Synthesis-Test-Analysis (DSTA) workflow.
Design: intuitive, data-driven, in silico modelling, docking etc…
Synthesis: can be linear or branched, which is preferred for late-stage functionalisation
Test: in vitro vs in vivo
Analysis: lots of data
Each iteration gives information on the toxophore and helps to build the Structure-Activity-Relationship (SAR).
Essentially, it is a multiparameter optimization algorithm.