Genomics, Epigenetics & Synthetic Biology

Part II Plant Sciences Module L1

Jim Haseloff
https://haseloff.plantsci.cam.ac.uk

Synthetic Biology and Plant Biotechnology

Lecture 1: Genetic modification and Synthetic Biology.
Lecture 2: Engineered DNA circuits.
Lecture 3: Reprogramming of multicellular systems.

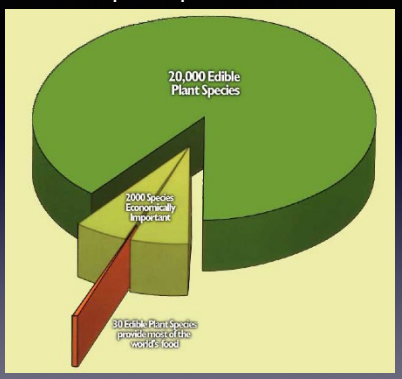
Lecture 1

1. **Origins of modern crops**
2. **Selection and breeding of new crop varieties**
3. **Genetic modification (GM) for plant improvement**
4. **Consolidation of agricultural biotech**
5. **From science to engineering**
6. **DNA technology**
7. **From Science to Engineering**
8. **Industrial revolutions**
9. **How to engineer biology?**

Lecture 2: Engineered DNA circuits.
Lecture 3: Reprogramming of multicellular systems.

Outline of this lecture

~400,000 plant species

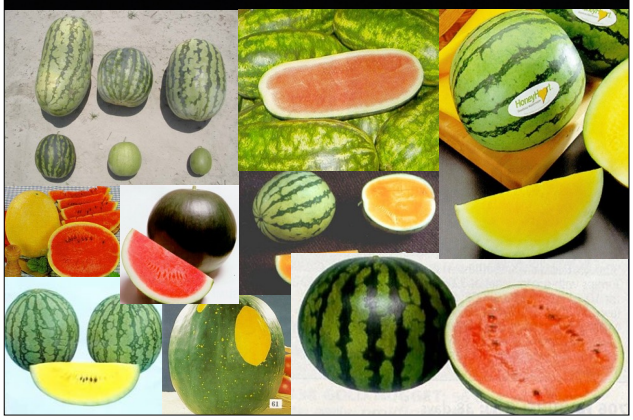


3 crop species (rice, wheat and maize) provide 60% of all calories and 54% of all protein in human food

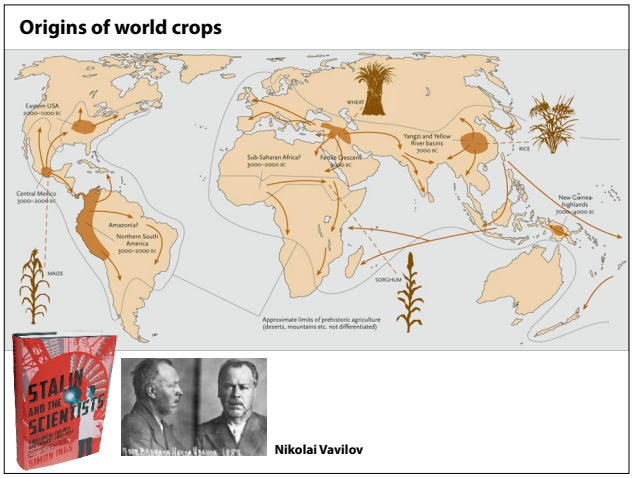
120 cultivated plant species

Crop plants sample a tiny fraction of total plant diversity. It estimated that there are around 400,000 plant species on Earth. Only around 20,000 of these have ever been used by humans as food, and only 2000 plant species have any economic importance as food crops. 30 species provide most of the world's food. Three species - rice, wheat and maize, provide 60% of calories and over half of the protein in human food. A vast reservoir of biological diversity remains untapped.

Plants are programmable



A wide variety of modern watermelon cultivars are shown. Selective breeding has allowed manipulation of the genetic content of the species to give rise to a range of useful crop traits - useful in agronomy and attractive to the consumer. The story of crops like the watermelon illustrate the flexible nature of plant development and growth, and the ability of humans to harness this. It provides an insight into what might be possible for many or all plant species - with the current adoption of faster, more powerful strategies for reprogramming plant genomes.



Nicolai Vavilov was a Russian biologist who first popularised the idea of geographical centres of diversity for the origin of modern crop species. These centres corresponded to areas of botanical diversity that coincided with the establishment of early human societies and plant domestication.

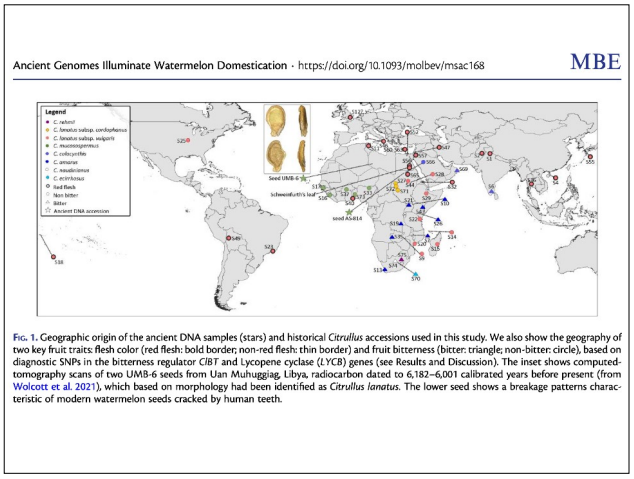
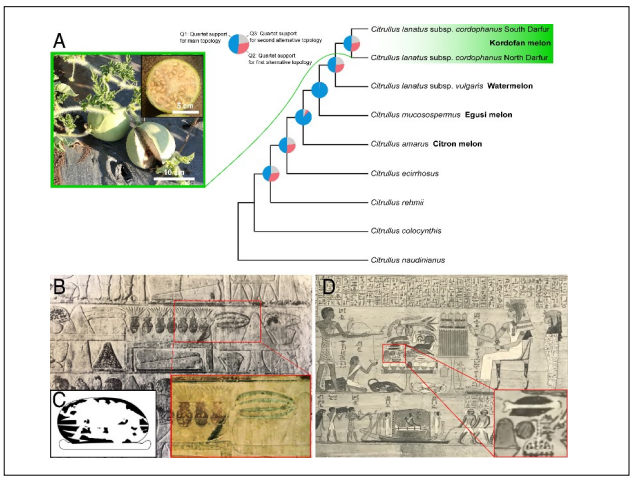
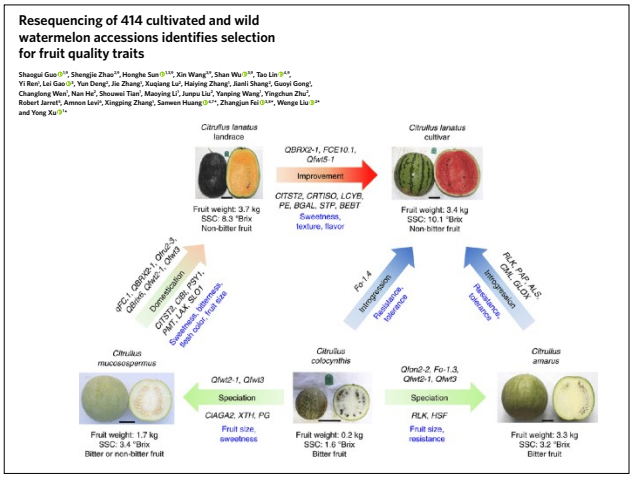


Fig. 1. Geographic origin of the ancient DNA samples (stars) and historical *Citrullus* accessions used in this study. We also show the geography of two key fruit traits: flesh color (red flesh: bold border; non-red flesh: thin border) and fruit bitterness (bitter: triangle; non-bitter: circle), based on diagnostic SNPs in the bitterness regulator *ClBT* and Lycopene cyclase (*LYCB*) genes (see Results and Discussion). The inset shows computed-tomography scans of two UMB-6 seeds from Uan Muhuggiag, Libya, radiocarbon dated to 6,182–6,001 calibrated years before present (from Wolcott et al. 2021), which based on morphology had been identified as *Citrullus lanatus*. The lower seed shows a breakage patterns characteristic of modern watermelon seeds cracked by human teeth.

The origin of domesticated watermelon was unclear until recently. Early collections (late 1700's) of presumed ancestors were mis-identified as Type Specimens by 20th Century taxonomists, and modern watermelons were thought to has originated in South Africa. However, the genome sequencing of the extant 7 Citrullus species, along with ancient preserved seeds (up to 6000 years old) suggests that the domestication took place in North East Africa, and the early crop was adopted in the Nile Valley.



Wild relatives or progenitors of crops are important resources for breeding and for understanding domestication. Identifying them, however, is difficult because of extinction, hybridization, and the challenge of distinguishing them from feral forms. Researchers have used collection-based systematics, iconography, and resequenced accessions of Citrullus lanatus and other species of Citrullus to search for the potential progenitor of the domesticated watermelon. A Sudanese form with nonbitter whitish pulp, known as the Kordofan melon (C. lanatus subsp. cordophanus), appears to be the closest relative of domesticated watermelons and a possible progenitor. These early forms may have been consumed primarily for the seed content, as the seed can be easily harvested and stored as a foodstuff. Recognisable images are seen in Egyptian tomb paintings that suggest that the watermelon may have been consumed in the Nile Valley as a dessert by 4360 BP. The genetic signature of bitterness loss is present in the Kordofan melon genome, but the red fruit flesh colour only became fixed in the domesticated watermelon. Mapping the genome variations in over 400 Citrullus accessions revealed shifts in allelic frequencies, suggesting that fruit sweetness has gradually increased over the course of watermelon domestication. That a likely progenitor of the watermelon still exists in Sudan has implications for targeted modern breeding efforts.



Population genomic analyses revealed better the evolutionary history of *Citrullus*, and suggested independent evolution and genetic exchange between *Citrullus amarus* and the lineage containing *Citrullus lanatus* and *Citrullus mucosospermus*. This indicated that different loci affecting watermelon fruit size have been under selection during speciation, domestication and improvement. A non-bitter allele, arising in the progenitor of sweet watermelon, is largely fixed in *C. lanatus*. Selection for flesh sweetness started in the progenitor of *C. lanatus* and continues through modern breeding - selecting for loci controlling raffinose catabolism and sugar transport. Fruit flesh colouration and sugar accumulation might have co-evolved through shared genetic components including a sugar transporter gene.

Archaeobotanical evidence provides a record of the origin and progress of domestication for the dessert watermelon. watermelons were domesticated for water and food there over 4000 years ago, and that sweet dessert watermelons emerged in Mediterranean lands by approximately 2000 years ago.
Descriptions in art and literature are consistent with evidence from genome sequencing and show the approximate timeline for spread and wider use of the crop - through to modern breeding and agriculture.

Watermelon phenotypes: ranging from the ancestral form (lower left), through to modern varieties. Modern breeding has produced an expanded variety of different characters including fruit colour, size and seed content. This illustrates a general feature of crop breeding. Plant form and secondary metabolism is highly plastic. Relatively few plant species have been domesticated and subjected to extensive breeding programmes. However for these few, it has been possible to generate very different cultivars adapted for human use, that breed true and grow vigorously, and form the basis of global agriculture for food production, animal feed and much bioproduction of materials.

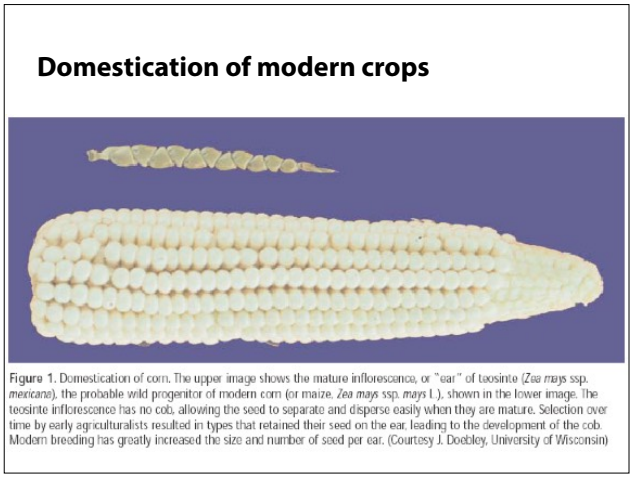
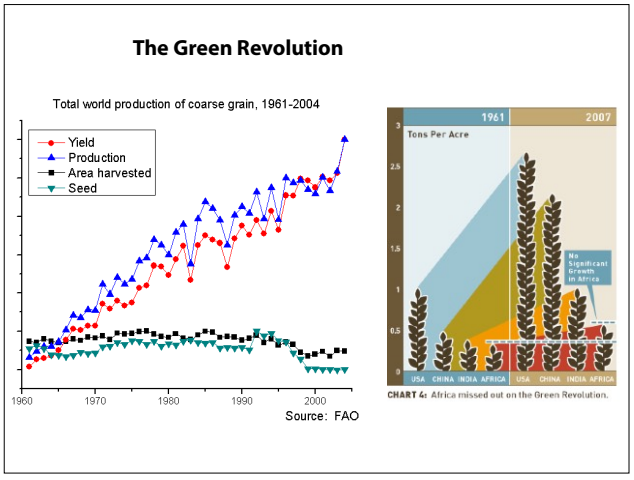


Figure 1. Domestication of corn. The upper image shows the mature inflorescence, or "ear" of teosinte (*Zea mays* ssp. *mexicana*), the probable wild progenitor of modern corn (or maize, *Zea mays* ssp. *mays* L.), shown in the lower image. The teosinte inflorescence has no cob, allowing the seed to separate and disperse easily when they are mature. Selection over time by early agriculturalists resulted in types that retained their seed on the ear, leading to the development of the cob. Modern breeding has greatly increased the size and number of seed per ear. (Courtesy J. Doebley, University of Wisconsin)

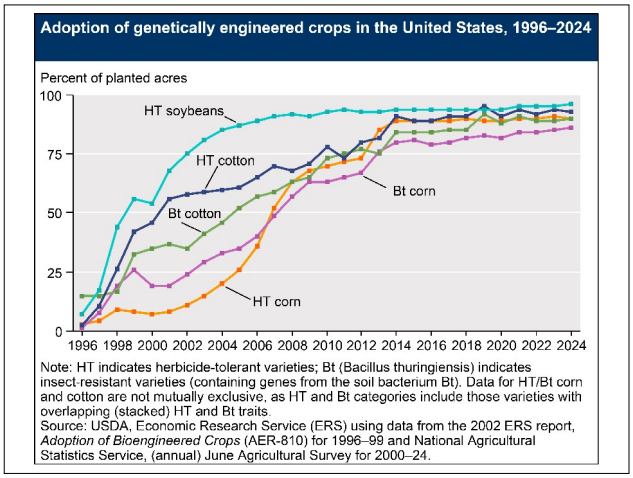
Early forms of maize strongly resemble teosinte, a plant endemic to Mesoamerica, and a subspecies of *Zea mays*. This likely progenitor has a strikingly distinct morphology, with smaller numbers of kernels arranged on a spike. It has been estimated that new varieties of maize been selected for over 9000 years. Modern varieties are characterised by a cob architecture with much larger numbers of kernels on each inflorescence. Europeans adopted maize as a crop and the 1800s saw large plantings across the Midwest of the United States. Before 1900 farmers in the Midwest were highly self-sufficient. They looked to the outside world for things like salt and nails, but external inputs into crops were minimal. Fertiliser inputs were limited to manure, pesticides were unknown and crops were true breeding and seed corn was obtained from previous year's crop. In the 1900s scientists

like G.H. Shull observed that open pollinated inbred forms of maize became less productive over time. In contrast heterosis or out-crossing gave rise to highly productive progeny. (Maize plants have separate male and female flowers and detasseling of male flowers is a simple way of ensuring selective crossing). Through the 1920s, plant breeding stations were established to create parental inbred lines that could be used for different crosses and to create highly productive maize seed. Entrepreneurs like Roswell Garst helped transform US agriculture last century. He helped to establish sales of hybrid corn seed with the noted corn breeder Henry Wallace in 1930s in Iowa. Wallace established Pioneer Hi-Bred, and Garst established Garst seed. There was a loss of diversity, from 786 varieties in 1903 to 52 in 1983. Farmers were previously highly self reliant - but the 20th

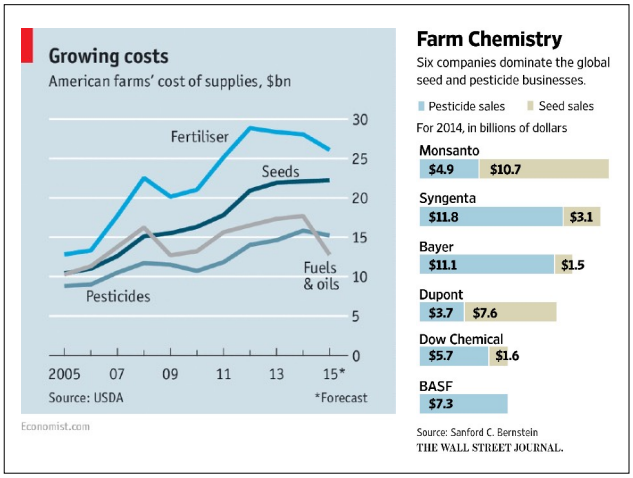
Selective breeding of other crops has dramatically improved their yields also. The decades following 1960's saw the breeding of highly productive new varieties of wheat. Many of these varieties were dwarf, which provided agronomic benefits and allowed commitment of more resources to seed production during growth. In addition, improved response to inorganic fertilisers and introduction of disease resistance through cycles of out-crossing and back-crossing contributed to new elite varieties. Shown above: "The harvesters" by Pieter Bruegel the Elder (1565) - with a graphic representation of a partly harvested wheat field in northern Europe. Note that the height of these wheat crops reached shoulder height. Modern wheat crops are much shorter, shown here with Norman Borlaug and colleagues at a trial field of Sonora-64. The story of Borlaug career is inspiring, a short version can be found at https://en.wikipedia.org/wiki/Norman_Borlaug. He has been credited with saving a billion people from starvation, and his work was extended to rice varieties.



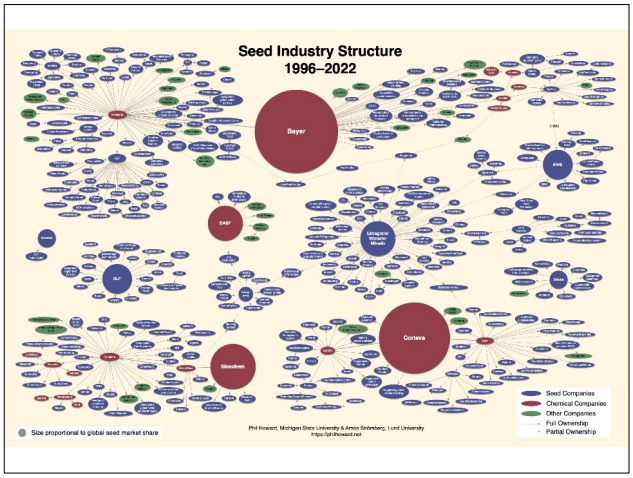
From the 1960s, the worldwide production of grain has increased dramatically in yield and total production despite relatively constant area of cultivation and planted seed. The bulk of these increases have been seen in the developed world, China and India. The benefits of increased production have not been so widely seen in Africa.



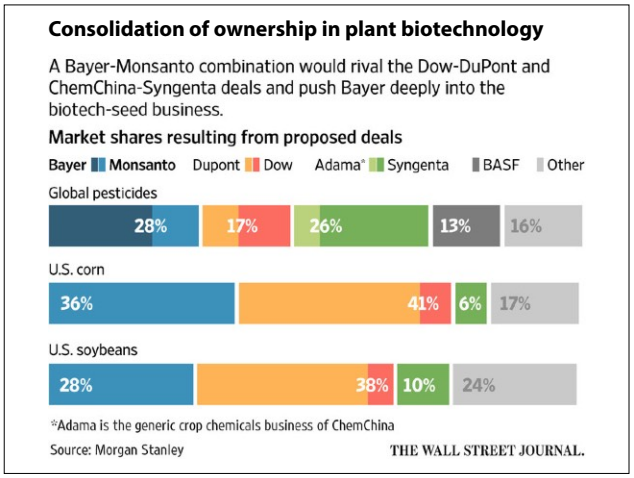
Until the early 1980s, the genetic modification of crops required the introduction of new genes through sexual crossing and refinement of traits through breeding. Specialised breeding techniques can allow access to gene pools outside of the same species - but access is confined to closely related plants. The advent of techniques to create transgenic plants allows synthesis of effectively any engineered DNA construct and unconstrained modification of plant genomes. This breakthrough came in 1983 with the independent publication of the first Agrobacterium-mediated plant transformation papers from three groups. The most predominant transgenic traits are herbicide and pest resistance. Countries in North and South America have seen the fastest and greatest increase in planting of biotech crops. They account for the overwhelming majority of GM producers globally. Outside of the Americas, there has been poor uptake of transgenic crops for food production. However, transgenic cotton is finding some adoption in Asia. Notably, there has not been wide adoption of transgenic crops in Europe or Africa to date.



The intensive nature of modern agriculture has led to increasing costs and complexity for farmers. Increasing yields come at the expense of increased fertiliser, pesticide, fuel and seed costs. The industry seen ever increasing levels of integration, so that a few companies are the major players in global agriculture.



The concentration of the Agricultural Biotech and Seed industry began in the 1970s and 1980s when intellectual property rights for seed improvements expanded, incentivising the research and development of new biotechnology seed traits and varieties by private companies. Improving biotechnology resulted in the first genetically modified crop varieties. Meanwhile, mergers between companies that specialised in pesticides, seed treatments, crop seeds and seed traits have over time resulted in an increasingly consolidated and highly integrated industry. There has been ongoing consolidation and concentration of the U.S. agriculture industry in recent decades. In 2015, six firms controlled the majority of the U.S. crop seed and agricultural chemical markets. Today, just four firms — Bayer, Corteva, ChemChina's Syngenta Group, and BASF — dominate those markets.



Six major agrochemical companies underwent mergers 2017-2019.

Disruptive technologies

Synthetic Biology: adoption of formal engineering principles in biology

The last few years have seen the emergence of new technologies for new engineering approaches that promise both highly efficient modular construction of DNA systems and systems for rational design. These have the potential to disrupt existing products and ways of working.

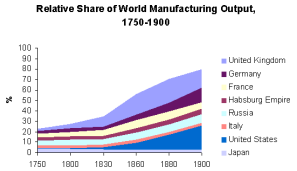
The Industrial Revolution:
based on innovations in coal, iron, steam and mechanical engineering

The late 1700s to early 1800's saw the emergence of new technologies and understanding of physical power and how this might be harvested and utilised by mechanical devices. (Here represented by Stevenson's "Rocket" the innovative forerunner of railways and global transport systems.)

First phase of the Industrial Revolution: innovation

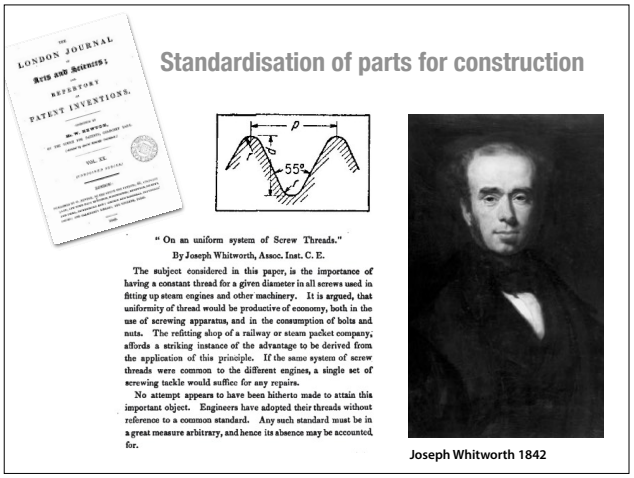
- **Steam power** - Improved steam engines were initially used for pumping out mines, but from the 1780s were applied to power machines. This enabled rapid development of efficient semi-automated factories
- **Iron founding** - Coke replaced charcoal in iron smelting. Improved production of bar iron, and eventually steel, resulted.
- **Textiles** - Cotton spinning was revolutionised by the invention of Richard Arkwright's water frame, James Hargreaves's Spinning Jenny, and Samuel Crompton's Spinning Mule). Similar technology was applied to spinning worsted yarn for various textiles and flax for linen.

Second phase of the Industrial Revolution: manufacturing

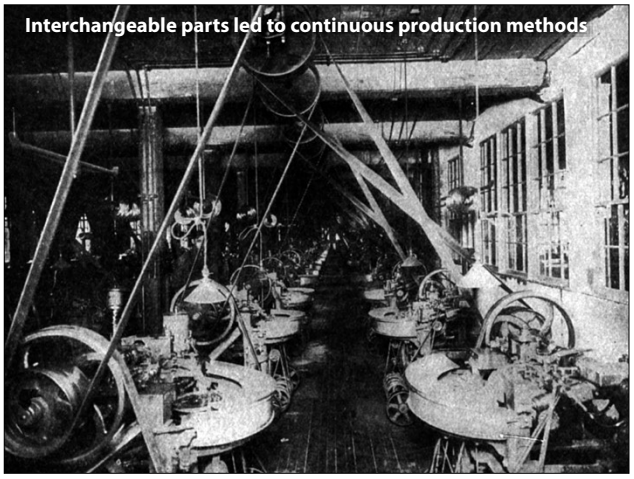


(Paul Bairoch, "International Industrialization Levels from 1750 to 1980")

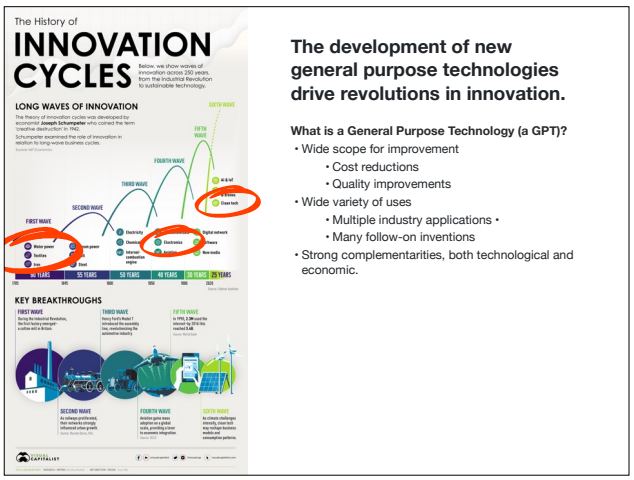
The Industrial Revolution progressed in stages. At first, raw technical developments, which stemmed from the late 1700's in Great Britain, had minimal impact on manufacturing output. At this point, the assembly of mechanical devices was bespoke. The design, assembly and interconnection of increasingly large and complex systems was difficult. What was required, was the development of new engineering standards that would allow simplification of these processes.



A simple example is the development of standardised screw threads for mechanical fasteners. Joseph Whitworth was awarded a patent for his establishment of a universal standard. Before this, individual machine shops would be machining incompatible fasteners. With the adoption of such standards, designers were liberated from these kinds of underlying details, and specialist manufacturers could emerge to supply a growing industry with standardised parts that everyone used.



The emergence of standardised parts and protocols for device interaction was required for the kinds of large scale machinery and power that was required for the establishment of continuous production methods, seen in the mid to late 1800's. It was this that led to large increases in manufacturing output, with economies of scale, and a synergistic effect on further development of the technologies.



The five waves that have crashed into our society so far are as follows: "Early Mechanisation" (1770s to the 1830s), "Steam Power and Railways" (1830s to 1880s), "Electrical and Heavy Engineering" (1880s to 1930s), "Fordist Mass Production" (1930s to 1970s) and "Information and Communication" (1970s to 2010s). https://www.maize.io/cultural-factory/technological-waves/

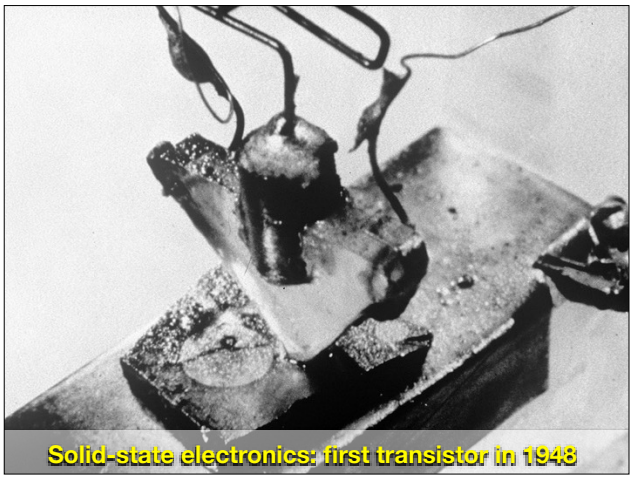
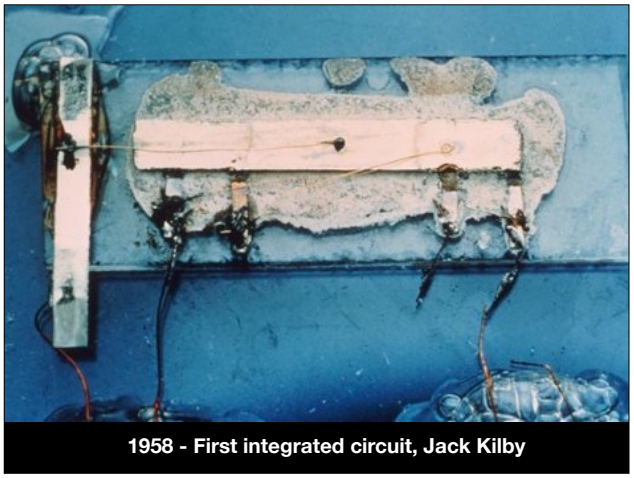
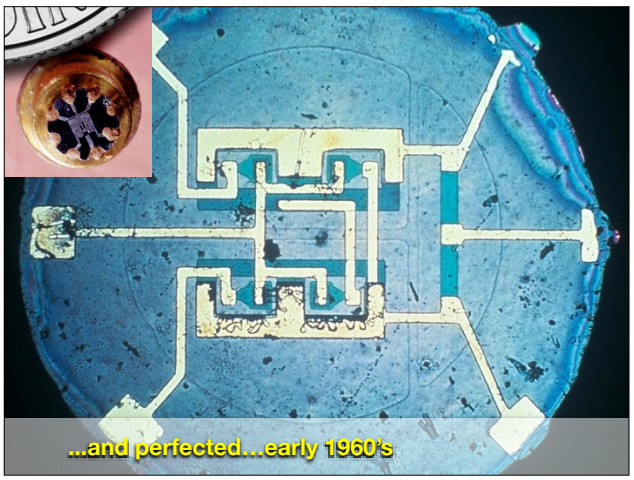


Photo of the first transistor, a bespoke device produced in 1948 at Bell labs. The simple solid-state device found immediate application as a low-voltage, low-power replacement for thermionic valves, used for electronic amplification and switching.

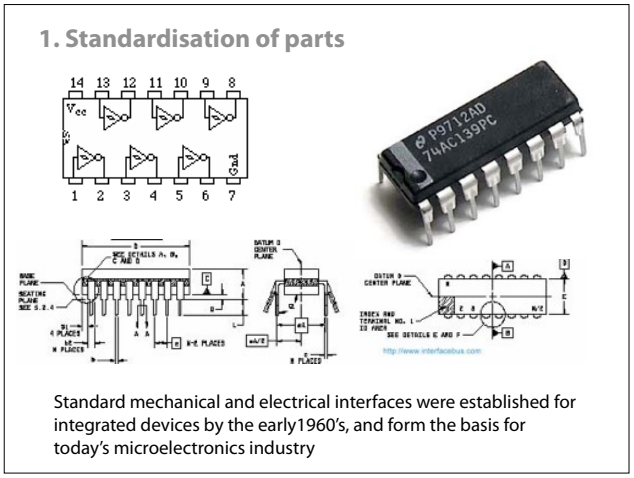
Within a few years, transistors had become commercial products and were sold widely. This generated additional development of the technology.



In 1958, Jack Kilby produced the first crude, handmade integrated circuit, which contained 5 logic devices on the same piece of semiconductor. This started a race to generate larger scale devices that lasts to this day.

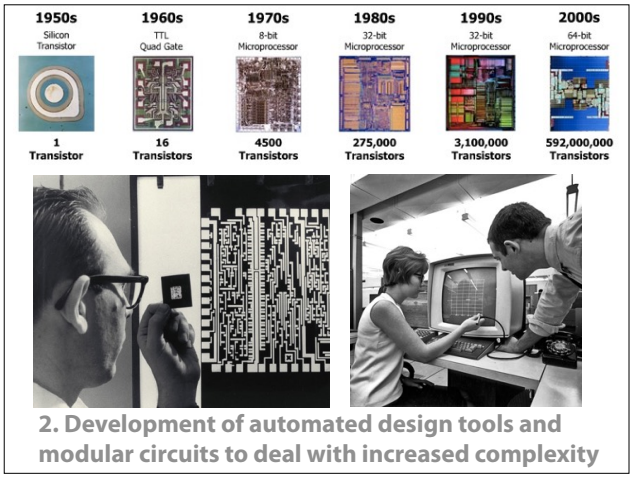


Within a few years the combination of new photolithographic and planar transistor techniques had created recognisable prototypes of the commercial devices that we would recognise today.

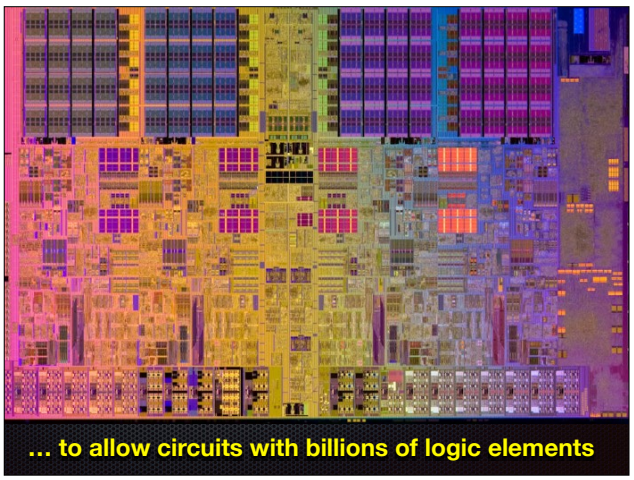


Within a few years mechanical and electrical interfaces had been standardised. This allowed the interoperability of these devices, and for engineers to mix devices from different sources.

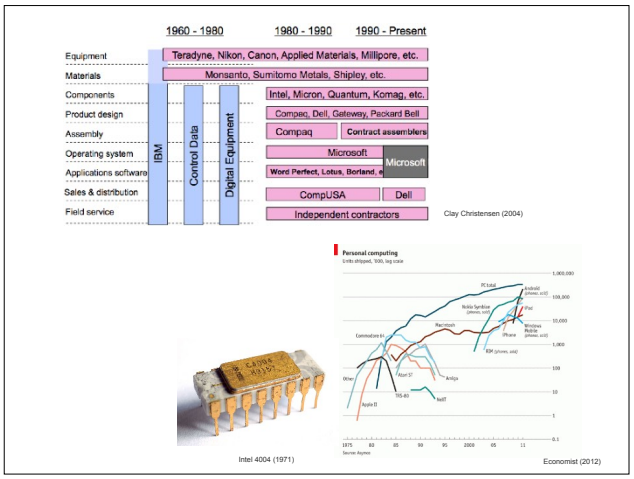
At first, these devices were designed by hand.



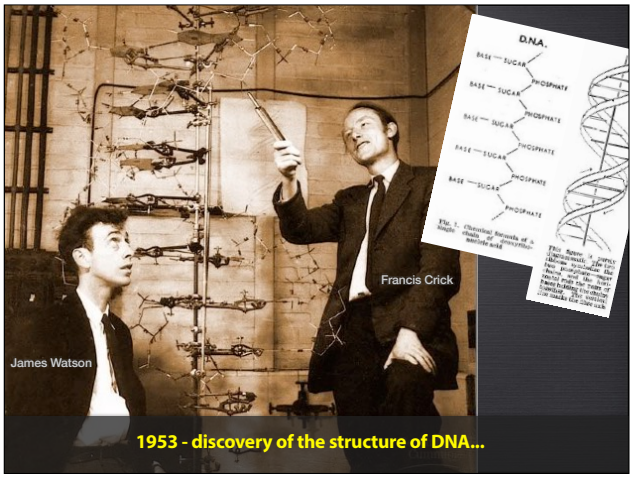
Increasing complexity saw the emergence of new automated design tools and reusable modular elements. Modularisation and standardisation are the hallmarks of modern engineering. They allow management of highly complex systems.



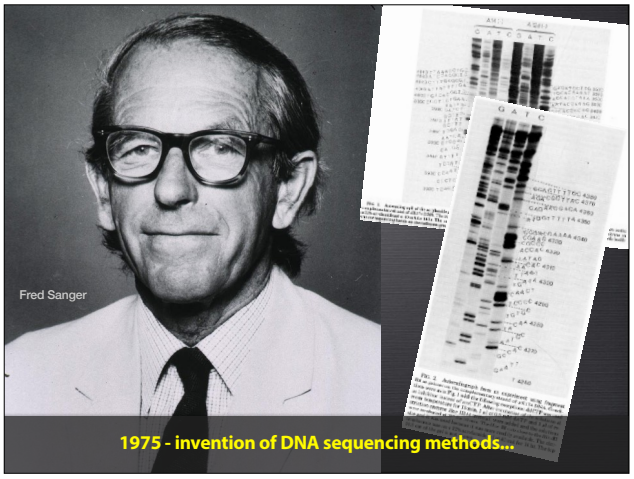
Modern integrated circuits contain billions of logic elements and themselves rely heavily on computer technology for their own design, testing and manufacture.
(e.g. M4 Max contains about 100 Billion transistors on a single chip with 3nm scale features)



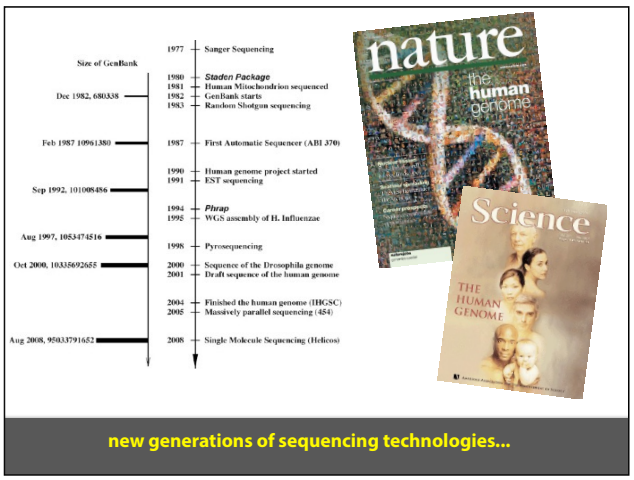
This level of consolidation has been seen in other industries. For example, the minicomputer industry was dominated by three companies (IBM, Control Data and DEC) through the 1960s. However the invention of the microprocessor in the early 70s, and the emergence of low-cost microcomputers caused disruption and saw the decline of these companies, and the emergence of a whole new range of businesses. The microcomputer industry was itself disrupted by the emergence of smart phones and apps. GM agribusiness is based on the use of 1980s technologies. Could this be due for disruption?



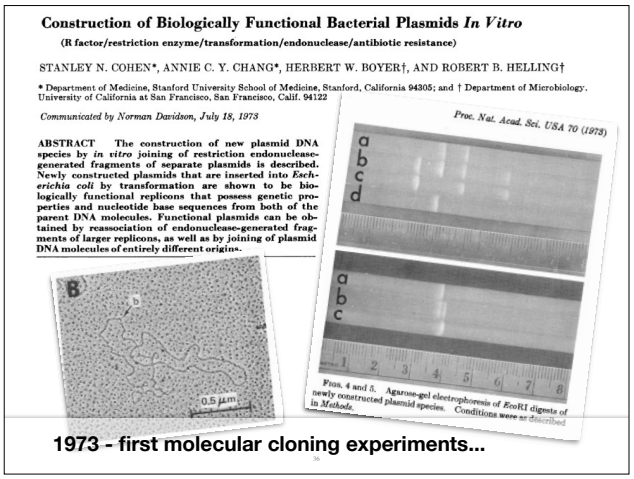
Over roughly the same time period, we have seen basic innovations in biology that allow similar engineering approaches. From discovery of the structure of DNA in 1953…



… To the development of DNA sequencing methods - at the kilobase-scale with Sanger sequencing in 1975



… Through to today's next generation gigabase-scale sequencing efforts.



The first molecular cloning experiments were published in 1973. In these first experiments DNAs were cut with restriction endonucleases, separated by electrophoresis, and pasted together with T4 DNA ligase. These experiments have triggered decades of genetic engineering experiments.

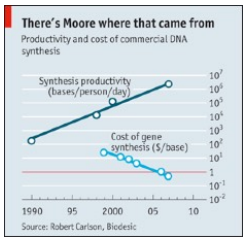
But...bespoke DNA assembly techniques are still common practice in the field after 40 years

Smithsonian museum exhibit: Stanley Cohen's laboratory bench

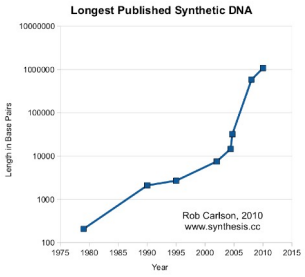
Photograph of the reconstruction Stanley Cohen's laboratory bench in the Smithsonian Museum. Not dissimilar to a modern molecular biologist's bench.

Enabling DNA Technologies

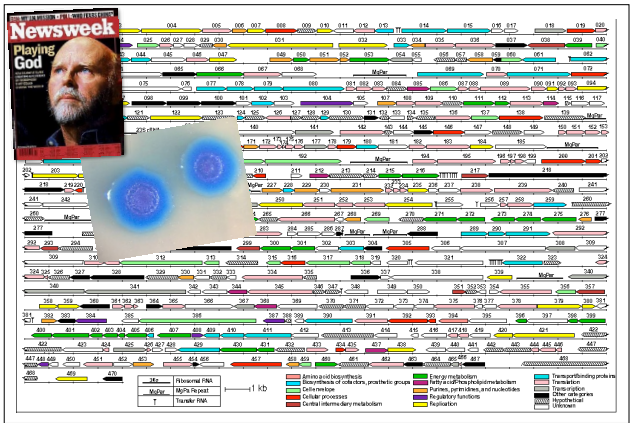
DNA Sequencing, DNA Assembly, DNA Synthesis &
Computational Modelling & Genetic Design



'Synthetic Biology - Life 2.0'
The Economist, August 31st 2006



Not only has the speed of DNA reading improved at an exponential rate, but the technology of DNA synthesis has also improved. It is now possible to synthesise DNA at dramatically lower prices, pennies per base-pair.



2008 - DNA synthesis of the first bacterial genome...

This has led to the construction of synthetic and refactored microbial genomes, for Mycoplasma (shown above), E. coli and yeast. From a technical point of view, we have the ability to generate unlimited new genomes, and potentially, generate novel species or variants. However (to date) all redesigned genomes have been forced closely follow the structures of natural systems. The current opportunity for synthesis is matched by the challenge of redesigning the DNA code for living systems, which possess complex, non-linear logic that is required for self-organisation, homeostasis and self-repair. Arguably the challenge of this century, and requires new engineering paradigms to deal with these levels of complexity. Synthetic Biology.

What is it?
Creation of artificial life?
Extension of genetic engineering?

syn·thet·ic [sin-**thet**-ik] –adjective

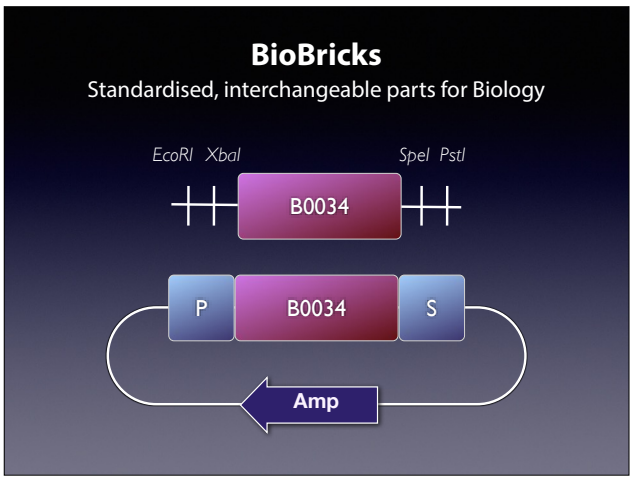
1. Prepared or made artificially, not of natural origin.

2. Relating to, or involving synthesis (construction of a coherent whole from separate elements)

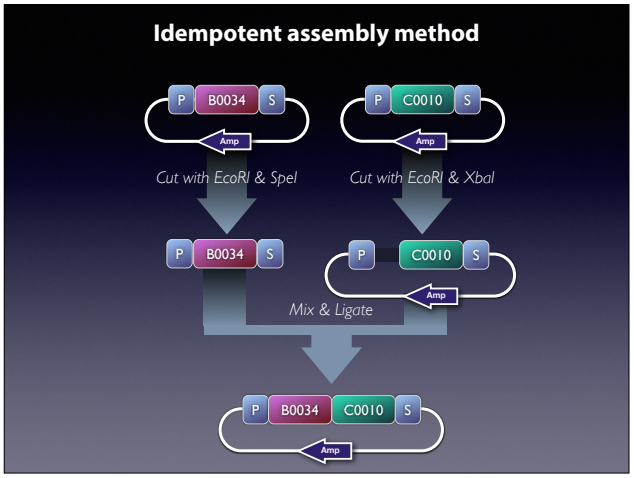
What is synthetic biology? The word "synthetic" can mean (i) artificial or (ii) relate to synthesis or construction. We will look at how the latter meaning can be used to refer to systematic approaches to biological construction.

Tom Knight is a computer science pioneer and the godfather of synthetic biology. After competing in the Science Talent Search programme in 1965 with an electroencephalograph (EEG) he built himself, Knight studied and taught at MIT. There, he worked in the emerging computer science field, developing the internet precursor, ARPANET. In the 1990s, Knight became interested in biology and leveraged his computer science background to create BioBricks, a type of DNA building block, and established the MIT Registry of Standard Biological Parts. In 2009, he co-founded Ginkgo Bioworks, which produces revolutionary organisms for commercial use. Ginkgo Bioworks went public in 2021 with the ticker symbol DNA. Knight is also co-founder of the international science competition iGEM. (https://golden.com/wiki/Tom_Knight_(scientist)-ZXDYE9)

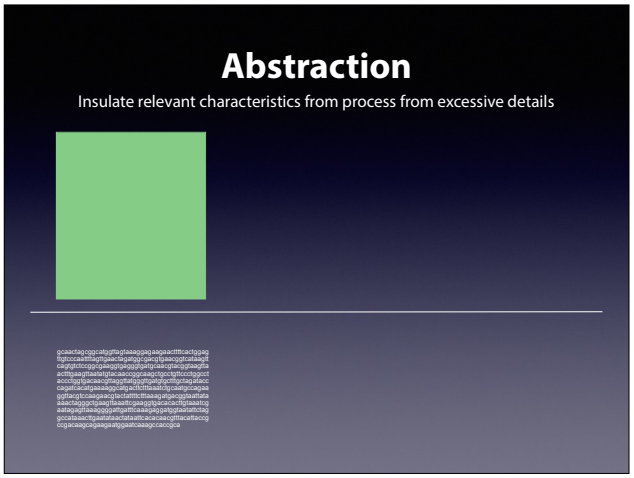
Tom Knight, a computer scientist at MIT, proposed a generalised method for large-scale DNA assemblies: **Idempotent Vector Design for Standard Assembly of Biobricks** (2003)
"The lack of standardization in assembly techniques for DNA sequences forces each DNA assembly reaction to be both an experimental tool for addressing the current research topic, and an experiment in and of itself. One of our goals is to replace this ad hoc experimental design with a set of standard and reliable engineering mechanisms to remove much of the tedium and surprise during assembly of genetic components into larger systems." http://hdl.handle.net/1721.1/21168



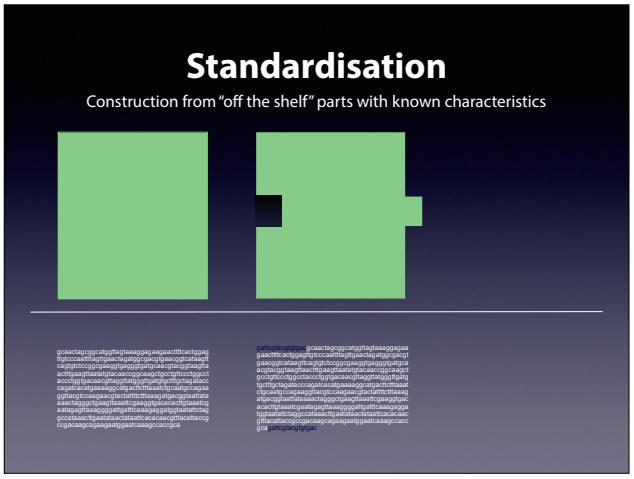
DNA parts would be composed in a standardised format for modular assembly. The modular parts would therefore be interchangeable, and…



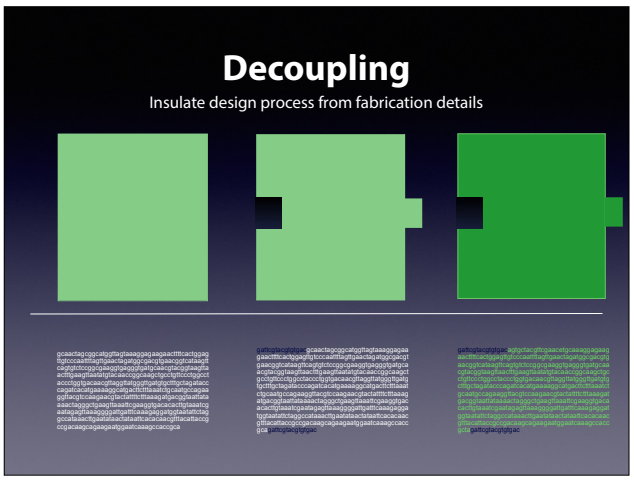
…the combination of any two parts would recreate the format of a standard part. (An object's properties remains unchanged during an idempotent operation). Note the arrangement of prefix (P) ands suffix (S) elements in this diagram, as two fragments are ligated.



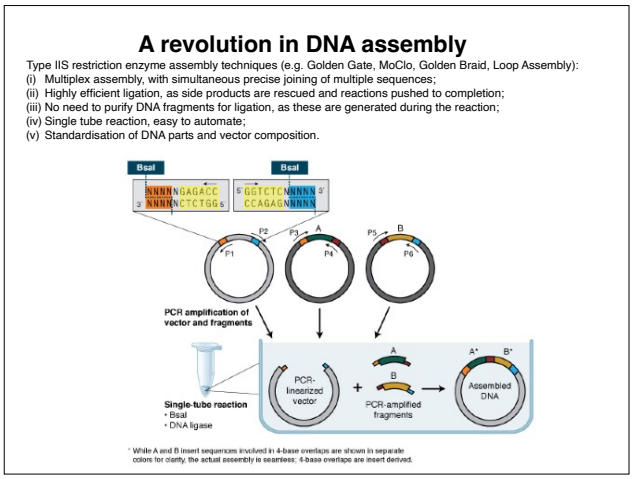
These advances have facilitated the efficient construction of engineered DNA sequences in a technical way. They also allow researchers to regard DNA encoded functions in a modular fashion. For example, this DNA part encodes the sequence of the green fluorescent protein. The modular nature of assembly standards can help insulate the designer from the underlying molecular-scale details of the DNA part.



Standardised parts come with an implied means of assembly. They can "plug" together in a manner similar to Lego parts.



The process of improvement of DNA parts can to be separated from the design process. In this case, a new part with a modified coding sequence for a brighter green fluorescent protein can be used interchangeably by a genetic circuit designer. The design process is decoupled from the fabrication process.

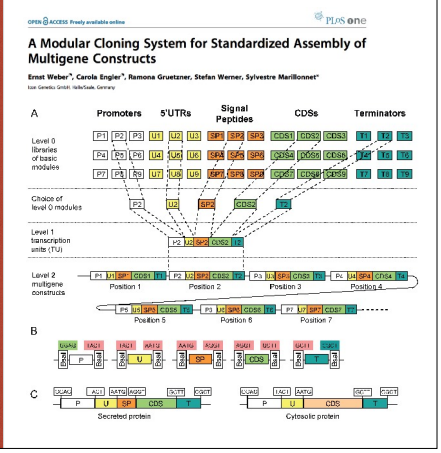


This approach has become increasingly sophisticated, now in the form of type IIS assembly techniques. These rely on restriction enzymes with cleavage sites that are offset from their recognition sequence. There is no need to isolate DNA fragments. Intact plasmid DNAs can be mixed, and cleavage and ligation of the fragments occurs in a single tube reaction to create the expected product.

Type IIS DNA assembly protocols:

Golden Gate
MoClo
ENSA
Golden Braid
Loop assembly:

adopted by the plant research community



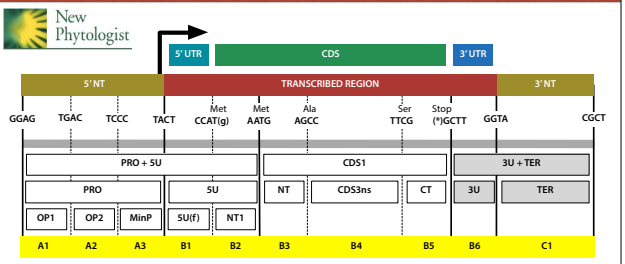
A Modular Cloning System for Standardized Assembly of Multigene Constructs

Ernst Weber*, Carola Engler*, Ramona Gruetzner, Stefan Werner, Sylvestre Marillonnet*

Type IIS assembly relies on the formatting of DNA fragments into particular classes. The different class fragments are then ligated to produce transcription units and can be further combined into a large multi-gene assemblies. The efficiency and ease of the assembly reactions has meant that this technique has been widely adopted by the plant research community.

A common syntax for plant DNA parts

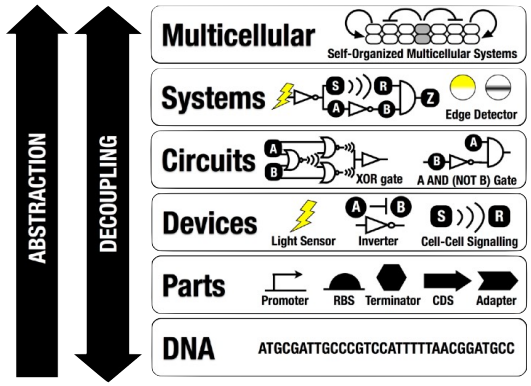
Based on Golden Gate standard assembly and type IIs restriction enzyme splints.



NPH-L-2015-19556.R1 Standards for Plant Synthetic Biology: A Common Syntax for Exchange of DNA Parts
by Patron, Nicola; Orzaez, Diego; Marillonnet, Sylvestre; Warzecha, Heribert; Matthewman, Colette; Youles, Mark; Raitskin, Oleg; Leveau, Aymeric; Farre-Martinez, Gemma; Rogers, Christian; Smith, Alison; Hibberd, Julian; Webb, Alex; Locke, James; Schornack, Sebastian; Ajioka, Jim; Baulcombe, David; Zipfel, Cyril; Kamoun, Sophien; Jones, Jonathan; Kuhn, Hannah; Robatzek, Silke; Van Esse, H Peter; Oldroyd, Giles; Sanders, Dale; Martin, Cathie; Field, Rob; O'Connor, Sarah; Fox, Samantha; Wulff, Brande; Miller, Ben; Breakspear, Andy; Radhakrishnan, Guru; Delaux, Pierre-Marc; Loque, Dominique ; Granell, Antonio; Tissier, Alain; Shih, Patrick; Brutnell, Thomas; Quick, Paul; Rischer, Heiko; Fraser, Paul; Aharoni, Asaph; Raines, Christine; South, Paul; And, Jean-Michel; Hamberger, Björn; Langdale, Jane; Stougaard, Jens; Bouwmeester, Harro; Udvardi, Michael; Murray, Jim; Ntoukakis, Vardis; Schafer, Patrick; Denby, Katherine; Edwards, Keith; Osbourn, Anne; Haseloff, Jim Haseloff

Further, plant researchers have adopted a common syntax for these plant parts to ensure interoperability across the community.

Hierarchical assembly



Federici, Rudge, Pollak, Haseloff, Gutierrez, 2013

The introduction of these engineering principles in biology is leading towards a more hierarchical way of constructing complex systems. DNA encoded functions can be formulated as standardised parts. These parts can be assembled into devices circuits and genetic systems - which can in turn be installed in multicellular systems.

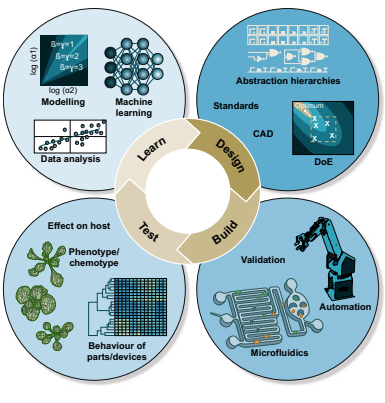


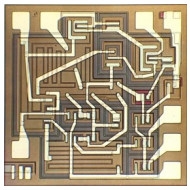
Fig. 1 Synthetic biology employs iterative cycles of design, build, test and learn. The data generated in each turn of the cycle are used to improve models of the system and inform the next cycle. Design is facilitated by the use of standards and abstraction hierarchies, Computer Aided Design (CAD) and the systematic application of statistics design of experiments (DoE), enabling large-scale experiments. Such experiments employ laboratory automation and microfluidics to increase reproducibility, minimise reaction volumes and reduce errors. Statistical and computational techniques are used to analyse data and refine predictive models.

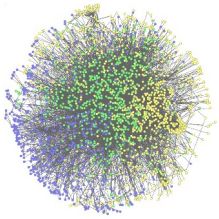
Nicola Patron, Tansley Review, New Phytologist

New tools for biological design, assembly, testing and learning come together in design–build–test–learn (DBTL) cycles, which are used iteratively in a wide variety of engineering disciplines - now implemented in genetic design.

The challenge of rewiring biological circuits

- Electronic circuit elements are physically insulated.
- Genetic circuits rely on molecular specificity within a cell.
- Cells provide insulation and additional scale of organisation.





Biological systems provide a new challenge for engineering. Even the most complex electronic device has constrained sets of interactions based on human design, and can be accurately modelled. Biological systems rely on molecular specificity, rather than defined connection paths for interactions. Cells contain complex compartments of interacting agents, with multiscale networks of substrates, enzymes and informational molecules, and interacting cells, the environment and other organisms. Biological systems have structures more like social networks, but where the genetic code and molecular specificity strictly regulate the behaviour of individuals, and the behaviour of populations are responsible for their self-organising properties.

Currently, the potential value of the new Synthetic Biology approach is recognised. To illustrate this: Tom Knight, with four of his ex-graduate students, from MIT formed a company (Ginkgo Bioworks). Ginkgo Bioworks terms itself the "organism design company", and provides a contract service with cutting edge tools for rebuilding biological pathways. The company issued an initial public offering in 2021. It raised over £15Bn. As a further illustration of the difficulty and long term nature of this challenge to reprogram natural systems, despite some successes, Ginkgo's stock has fallen from an initial high of several hundred dollars to around $7 today.

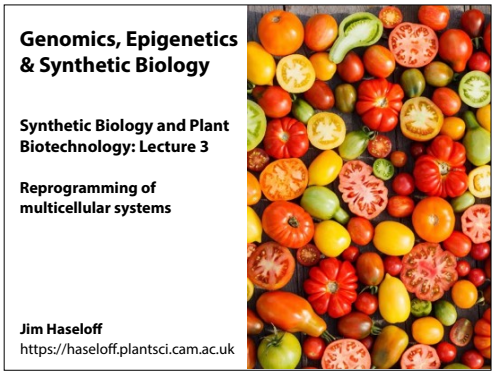
Lecture 1

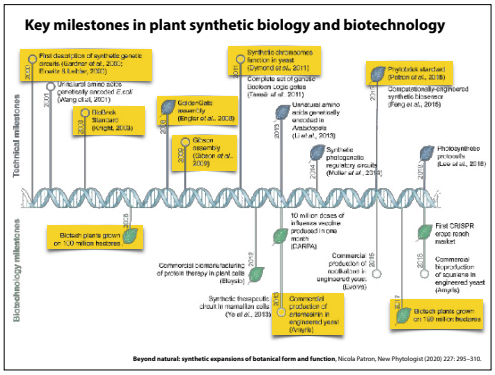
1. **Origins of modern crops**
2. **Selection and breeding of new crop varieties**
3. **Genetic modification (GM) for plant improvement**
4. **Consolidation of agricultural biotech**
5. **From science to engineering**
6. **DNA technology**
7. **Synthetic Biology**

Lecture 2: Engineered DNA circuits.

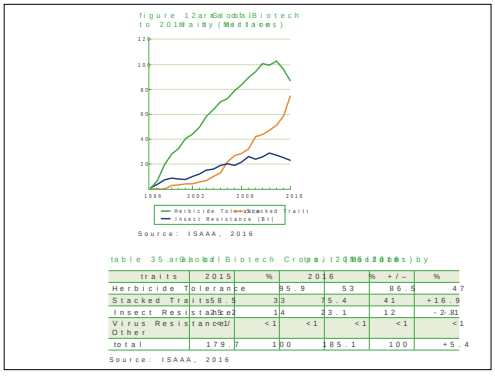
Lecture 3: Reprogramming of multicellular systems.

The next lecture will focus on engineering DNA circuitry.



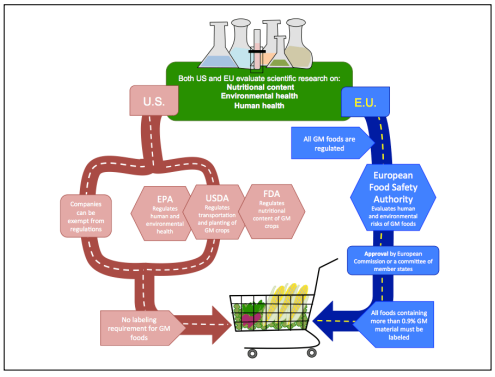


Key milestones in synthetic biology and biotechnology related to plant systems. Leaf symbols indicate milestones in plant and plant-related systems, and circles indicate related technical advances. From: Beyond natural: synthetic expansions of botanical form and function, Nicola Patron, New Phytologist (2020) 227: 295–310.



The first transgenic plants were created in the laboratory in the early 80s. By the mid-90s field trials of transgenic crops were underway. The first generation of traits included herbicide tolerance for weed control, and insect and virus pest resistance. In the subsequent 20 years there has been a rapid uptake in the use of these single gene traits in maize, cotton and soybean crops. We are seeing a sharp rise in the use of combined, or stacked, traits. In 2016, 185 million ha of transgenic crops were grown.

The introduction of unlabelled GM corn and soybean products from the US during the 90s caused a consumer backlash in Europe. This is partly due to the lack of choice and benefit for the consumer and perceived risks associated with the new technology - in the wake of the BSE crisis. Further there has been strong distrust of the large agrochemical companies who are exploiting the new technology.

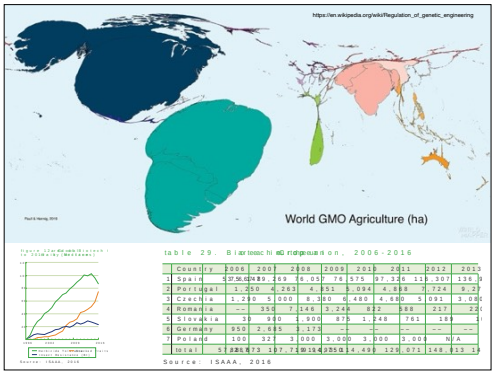


The US and Europe have adopted very different regulatory systems for GM foods. Food companies submit the same types of scientific data to U.S. and EU regulatory bodies for approval. Three separate agencies in the U.S. evaluate the potential risks of GM foods, while a centralised approval process is established in the EU. Approval and labeling requirements are stricter in the EU. (http://sitn.hms.harvard.edu/category/flash/special-edition-on-gmos/)

Different approaches to GMO regulation:

Precautionary Principle (Europe)- GM crops are potentially dangerous and pose new risks and thus their use should be

avoided until they are proven safe.
Substantial Equivalence Principle (USA) - GMOs are no different from conventional crops, if the products so derived are "substantially equivalent" in composition, nutritive value or safety after thorough comparative testing.

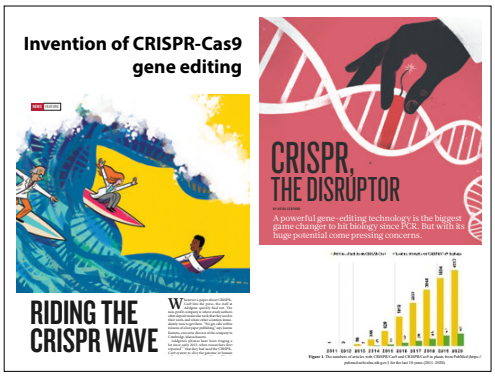


Despite the rapid adoptance of GM technology for crop improvement in different global markets, regulatory and consumer driven barriers to acceptance have limited uptake of the technology in Europe. (https://food.ec.europa.eu/plants/genetically-modified-organisms/gmo-legislation_en) Stricter regulatory limits on planting of GM crops came into effect in 2003, and while national exceptions were allowed in 2018, cropping of GM plants remains relatively low. A world map is shown, where national boundaries have been adjusted to reflect the relative areas of GM crops grown in each country (dated 2019).

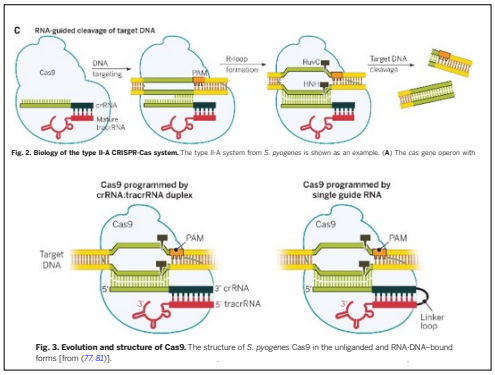
Disruptive technologies:

Genome editing

The last few years have seen the emergence of both new technologies for direct genome editing, and for new engineering approaches that promise both highly efficient modular construction of DNA systems and systems for rational design. These have the potential to disrupt existing products and ways of working.

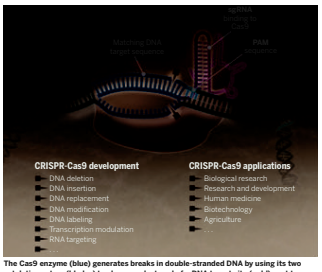


Following the development of plant transformation techniques in the early 1980's, and the first GM crops a decade later - the last decade has seen an explosion of new gene editing techniques and their application for biomedical and agricultural uses.



The CRISPR class of gene editing tools are derived from natural systems for bacterial immunity. Bacteria contain mechanisms for converting foreign DNA to embedded interspersed segments of sequence of defined length - the CRISPR arrays. These act as a reservoir of elements that can be used to attack incoming homologous sequences - such as phage DNAs. CRISPR sequences are transcribed, paired with the tracrRNA and bound to the Cas9 protein to produce a targeted, RNA-programmed nuclease.
The tracrRNA and crRNA components of the nuclease can be fused to create a single guide sequence that, in combination with Cas9, will produce a nuclease that can be targeted to

any DNA sequence adjacent to a 3 nucleotide PAM sequence.



The Cas9 enzyme (blue) generates breaks in double-stranded DNA by using its two catalytic centers (blades) to cleave each strand of a DNA target site (gold) next to a PAM sequence (red) and matching the 20-nucleotide sequence (orange) of the single guide RNA (sgRNA). The sgRNA includes a dual-RNA sequence derived from CRISPR RNA (light green) and a separate transcript (tracrRNA, dark green) that binds and stabilizes the Cas9 protein. Cas9-sgRNA–mediated DNA cleavage produces a blunt double-stranded break that triggers repair enzymes to disrupt or replace DNA sequences at or near the cleavage site. Catalytically inactive forms of Cas9 can also be used for programmable regulation of transcription and visualization of genomic loci.

The CRISPR-Cas9 system can be used to create a programmable DNA binding complex. This will normally create a double-strand break at the target site. This has ben used widely for targeted mutagenesis, via error-prone repair of dsDNA breaks in vivo, and to promote DNA replacement through homologous repair. In addition, the CRISPR-Cas9 complex has been engineering to have a wide range of other activities…

The ribonucucleoprotein is finding many applications as a tool for genome manipulation *in vivo* - new functions can be genetically fused to the Cas9 protein and used to deliver the activity to precise locations on a DNA molecule. This can apply not just for DNA cleavage and editing, but also selective base editing, regulation of transcription, chemical modification of chromatin, histone modification, etc.

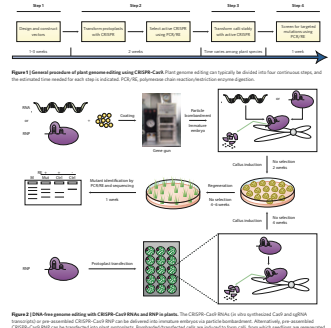


Figure 1 | General procedure of plant genome editing using CRISPR-Cas9. Plant genome editing can typically be divided into four continuous steps, and the estimated time needed for each step is indicated. PCR/RE, polymerase chain reaction/restriction enzyme digestion.

Figure 2 | DNA-free genome editing with CRISPR-Cas9 RNAs and RNP in plants. The CRISPR-Cas9 RNAs (in vitro synthesized Cas9 and sgRNA transcripts) or pre-assembled CRISPR-Cas9 RNP can be delivered into immature embryos via particle bombardment. Alternatively, pre-assembled CRISPR-Cas9 RNP can be delivered into protoplasts, from which transgene-free mutant plants with induced heritable mutations are regenerated under the selection-free conditions. Regenerated plants are screened for mutations via PCR/RE assay and sequencing. The embryogenic callus method is also indicated for most steps. Delivering CRISPR-Cas9 reagents via RNP limits their temporal activity, thereby improving their precision. RE, restriction enzyme; M, DNA marker; mut, mutant; ctrl, control.

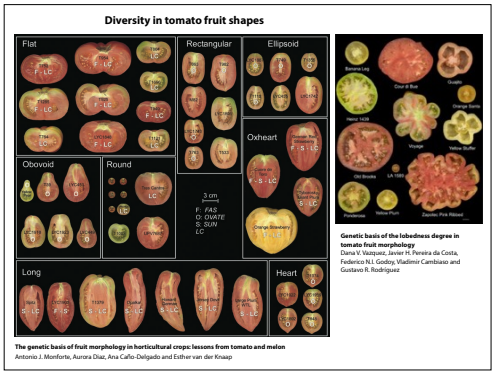
DNA-free manipulation of crop plants. Delivery of CRISPR-Cas9 ribonucleoprotein into plant cells by protoplast transformation or biolistic delivery allows precise manipulation of plant genomes without the introduction of plant pathogen sequences (e.g. Agrobacterium), or other foreign DNA. This allows the production of modified plants with engineered genomes - which would be indistinguishable from, say, mutant plants produced by random mutagenesis.

Fig. 1 Regulatory status of gene-edited crops (when no foreign DNA is inserted). Dark green = regulated as conventional crops. Pale green = draft regulations suggest they will be regulated as conventional crops. Red = viewed as GMOs. Yellow = under review but likely to be viewed favorably. The UK (shown in yellow) recently gave the go ahead for research field trials to proceed without the need for GMO regulatory oversight (when no foreign DNA is inserted); however, restrictions currently remain for commercial applications. Further amendments are still under review but look favorable

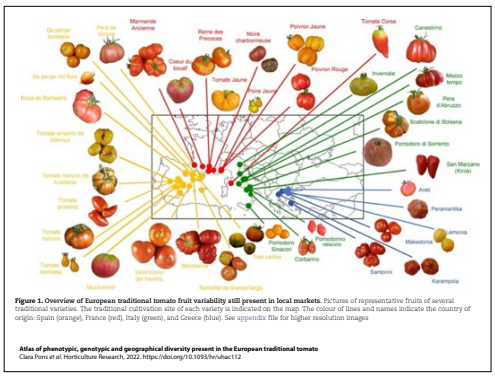
In many countries (green), gene-edited crops are regulated as conventional crops, provided that they don't contain foreign DNA. Gene-editing approaches have lowered the threshold for development and cultivation of new plant varieties, and provided a major incentive for new businesses. However, gene-edited plants are still regarded as GMOs in Europe and some other countries (red), subjecting their use to much higher levels of regulation.

Example: Altering traits in tomato

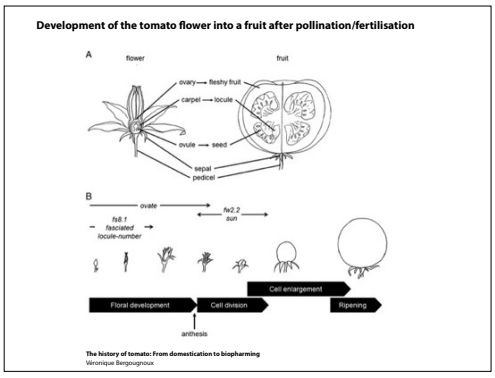
Targeting genes using CRISPR-mediated mutagenesis



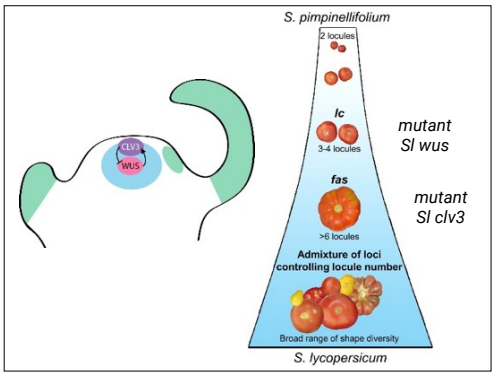
The domestication of a crop plant like the tomato, has been accompanied by the selection and breeding of a wide variety of variants. These include plant varieties with profound differences in fruit size and shape, and plant architecture. An atlas of different tomato fruits showing attempts to classify the wide variety of fruit shapes and architecture, including locule number. Images taken from: The genetic basis of fruit morphology in horticultural crops: lessons from tomato and melon. Antonio J. Monforte, Aurora Diaz, Ana Caño-Delgado and Esther van der Knaap, and Genetic basis of the lobedness degree in tomato fruit morphology, Dana V. Vazquez, Javier H. Pereira da Costa, Federico N.I. Godoy, Vladimir Cambiaso and Gustavo R. Rodríguez.



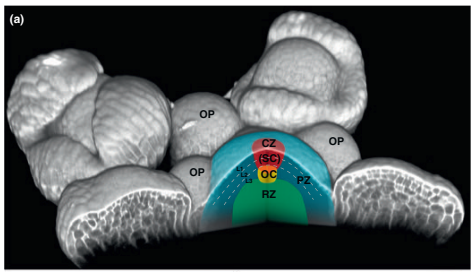
Modern high-throughput sequencing techniques have allowed determination of the entire genomes of modern and heritage varieties of tomatoes. For example, the introduction of tomatoes into Europe in the late 1400's was followed by extensive selection of distinct varieties across Spain, France, Italy and Greece. The origins and appearance of the fruit of a number of these varieties are shown on the slide. The genomes of these diverse varieties were sequenced and compared to help identify genetic variation associated with particular traits.



The core cellular architecture of the tomato fruit is laid down during floral development. The slide shows homologous domains in the flower and fruit, and an outline of the developmental timing that gives rise to the mature fruit. Alterations in early development can give rise to different morphological outcomes.



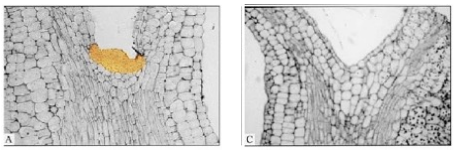
Modern tomato varieties emerged from wild Peruvian species (*Solanum pimpinellifolium*). Two mutant alleles which have played a major role in the breeding of large fruit sizes are *Locule number* (*lc*) and *Fasciated* (*fas*). The conserved Clavata3 (CLV3) - Wuschel (WUS) negative feedback circuit controls meristem and fruit size in plants. *Fasciated* encodes the *Solanum lycopersicum* CLV3 gene, and *locule number* (*Lc)* encodes the WUS gene. Mutations in the fasciated (fas) and locule number (*lc*) both contributed to increased tomato fruit size and locule number during domestication.



Engineering multicellular architectures in plants

In order to reprogram plant systems, it is necessary to regulate cell-cell communication. Meristematic zones initiate, maintain and coordinate patterns of gene expression in a multicellular context. They produce primordia that continue to grow and produce the organs of a plant. CRISPR-Cas9 tools are now being used to manipulate shoot growth in crop plants.

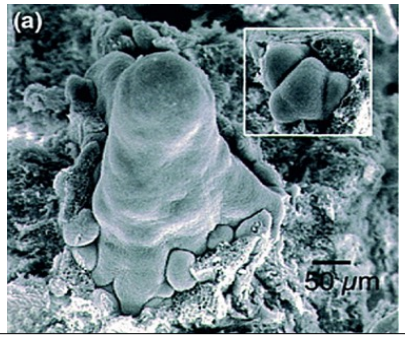
WUSCHEL (WUS) is a homeodomain gene that is required for formation and maintenance of the shoot apical meristem in Arabidopsis.



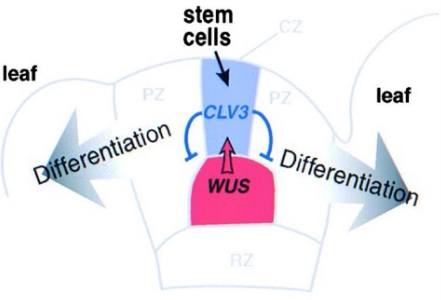
wt *wuschel (wus)*

The growth of the apical meristem in high apical meristem in higher plants is controlled by opposing sets of regulatory proteins. We will look at two of these, as characterised in the model plant Arabidopsis. First, *Wuschel* is a homeodomain gene that is required for maintenance of the stem cells within the shoot apical meristem. In a *wuschel* mutant, the cells within the apical meristem fully differentiate, and the meristem terminates. Overexpression of the *Wuschel* gene results in the hyper-proliferation of stem cells within the meristem. Wuschel is sufficient for the production of indeterminacy in targeted cells.

Clavata3 is a negative regulator of Wuschel gene activity



In contrast, the *Clavata3* gene plays a role in suppressing the size of the meristem. In its absence the meristem is enlarged (which can result in oversize organs). Overexpression of Clavata3 results in loss of meristem activity (and undersized flowers). Scanning electron micrographs of wild-type (inset) and *clavata3* mutant meristems.



The Stem Cell Population of *Arabidopsis* Shoot Meristems Is Maintained by a Regulatory Loop between the *CLAVATA* and *WUSCHEL* Genes
Heiko Schoof, Michael Lenhard, Achim Haecker, Klaus F.X Mayer, Gerd Jürgens and Thomas Laux

The opposing influences of the Wuschel (WUS) and Clavata3 (CLV3) genes contribute to a cellular feedback system that contributes to initiation and maintenance of meristem size and activity in plant apices. Expression of Wuschel drives stem cell proliferation, which in turn triggers higher levels of CLV3 production. The CLV3 gene product is a small, secreted peptide that diffuses across the meristem via the extracellular spaces between cells - to bind to receptor proteins and inhibit WUS activity. This regulatory loop plays a key role in the maintenance of meristem size. Perturbations of this regulatory loop result in changes in meristem size.

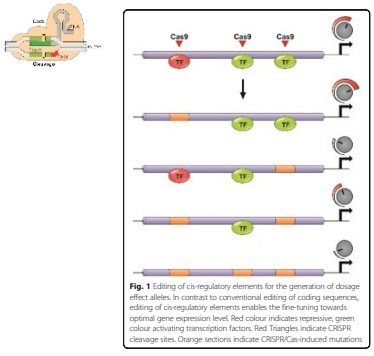
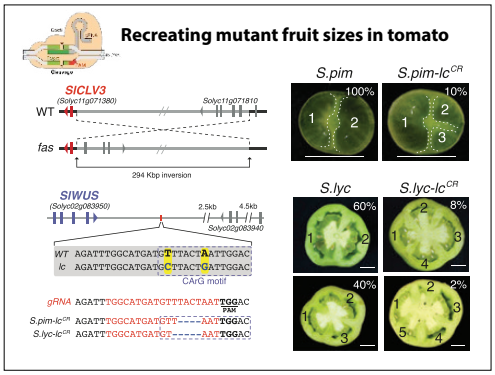


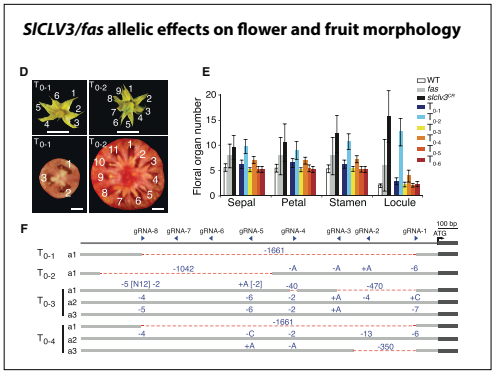
Fig. 1 Editing of cis-regulatory elements for the generation of dosage effect alleles. In contrast to conventional editing of coding sequences, editing of cis-regulatory elements enables the fine-tuning towards optimal gene expression level. Red colour indicates repressive, green colour activating transcription factors. Red Triangles indicate CRISPR cleavage sites. Orange sections indicate CRISPR/Cas-induced mutations



During domestication:

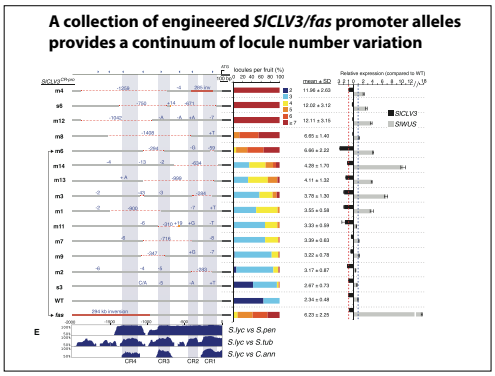
1. The fas mutation was caused by an inversion with a breakpoint 1 Kbp upstream of SlCLV3. (D)
2. The lc QTL (red rectangle) is associated with two SNPs (in bold) in a putative repressor motif (CArG, blue-dashed square) 1.7 Kbp downstream of SlWUS.

Mutations that affect the crop phenotype can be recreated and combined in different (or wild) varieties. CRISPR/Cas9-induced deletions in the CArG repressor motif are shown (blue-dashed square) in Solanum pimpinellifolium (S. pim) and Solanum lycopersicum (S. lyc). The gRNA target sequence is highlighted in red and the PAM site underlined. S.pim-lcCR plants produce fruits with more than two locules. S.pim-fasNIL S.pim-lcCR double mutants synergistically increase locule number.



DNA-free manipulation of crop plants. Delivery of CRISPR-Cas9 ribonucleoprotein into plant cells by protoplast transformation or biolistic delivery allows precise manipulation of plant genomes without the introduction of plant pathogen sequences (e.g. Agrobacterium), or other foreign DNA. This allows the production of modified plants with engineered genomes - which would be indistinguishable from, say, mutant plants produced by random mutagenesis.

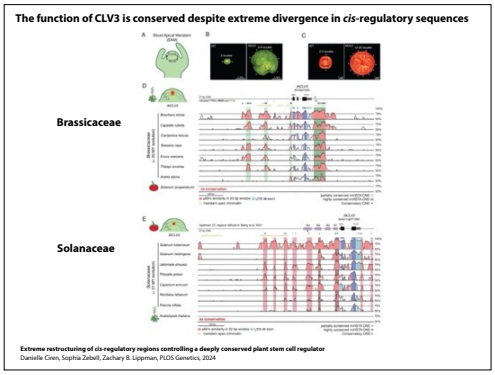
Schematic of SlCLV3 promoter targeted by eight gRNAs (numbered blue arrowheads). Blue arrows, PCR primers. Weak and strong effects on flower morphology and fruit size were observed among T0 lines. Number of floral organs and locules are indicated. (E) Quantification of floral organ number in T0, WT, fas, and slclv3CR plants. (F) Sequencing of SlCLV3 promoter alleles for all T0 plants. Deletions (–) and insertions (+) indicated by numbers or letters. T0-5 and T0-6 contained only WT alleles. Blue arrowheads, gRNAs; a, allele.



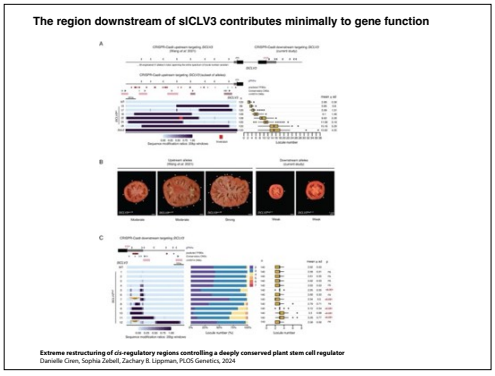
1. Sequences of 14 new SlCLV3pro alleles. Deletions (–) and insertions (+) indicated as numbers or letters. gRNAs, blue arrowheads. Parental F1s marked at right.
2. qRT-PCR of SlCLV3 and SlWUS from reproductive meristems for WT, fas, and each SlCLV3CR-pro allele. Dashed lines mark WT levels for SlCLV3 (red) and SlWUS (blue)

Example: dissecting genome functions

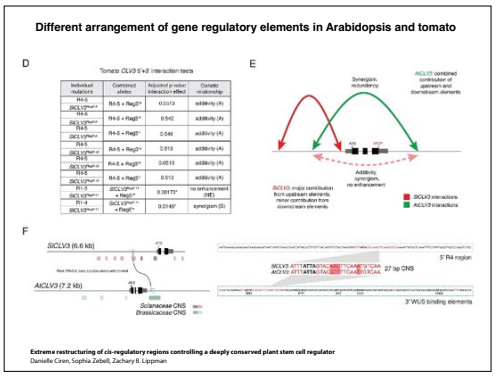
Discovering complex genetic mechanisms



The Lippman lab at CSHL investigated the gene regulatory functions of the Arabidopsis and tomato CLV3 genes, which diverged over 125 million years ago. Using *in vivo* genome editing, they generated over 70 deletion alleles in the upstream and downstream regions of the CLV3 genes and compared their individual and combined effects on the number of carpels that make fruits. Despite strong sequence conservation, they found remarkable differences in the cis-regulation of this deeply conserved plant stem cell regulator. The finding underscored the need for species-specific dissection of regulatory mechanisms to effectively engineer trait variation in different crops. (https://doi.org/10.1371/journal.pgen.1011174)

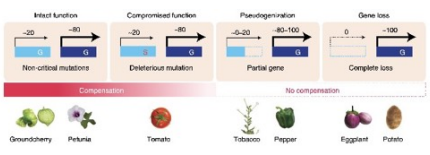


Most of the cis-regulatory functions that control the CLV3 gene in tomato reside in the 5′ proximal intergenic region. CRISPR-induced mutations in this region cause substantial changes in phenotype (B). These types of in vivo editing of gene functions have also unmasked cryptic forms of compensation for gene function.



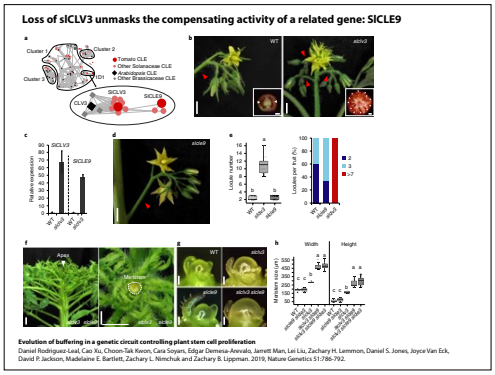
Rearrangement of gene regulatory sequences. A model summarising the relative contributions of the 5′ and 3′ region, as well as their interactions, to the regulation of *Arabidopsis thaliana* and tomato (*Solanum lycopersicum)* CLV3 genes. (F) A conserved 27 bp sequence which overlaps with the distal R4 region in the tomato 5′ (outlined by a purple dashed box), and a known WUS transcription factor binding sequence in the Arabidopsis 3′ (outlined by a blue dashed box). The DNA sequences within these regions are shown, with the 27 bp sequence in red, nucleotide mismatches highlighted in red, and the core ATTA WUS binding element in black. All five of the previously characterised AtCLV3 3′WUS binding elements are also shown in bold and named according to their position.

Unmasking circuit buffering:
Finding genetic compensation

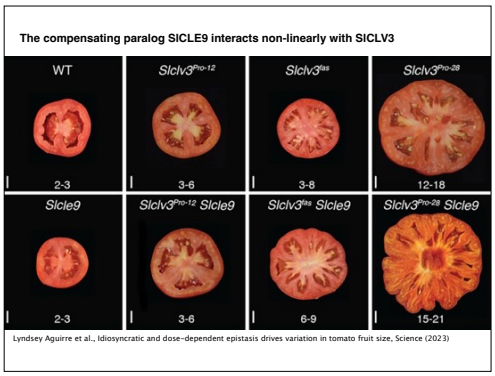


Variation in CLV3-related gene number and their expression across the Solanaceaceae plant family. Dark blue and blue rectangles indicate the coding region of CLV3-related genes. Arrows and their thickness represent gene expression and their relative levels. Numbers above the arrows indicate hypothetical relative proportions of orthologue expression levels. ‘G’ and ‘S’ within the rectangles denote the sixth amino acid of each CLV3-related dodecapeptide. Dashed rectangles mark deletions of the coding region, resulting in pseudogenes (pepper and tobacco) and complete gene loss (eggplant and potato) in each genome. Mutation of the CLV3 gene can trigger functional compensation by the closely related CLE9 gene in some members of the plant family. The red gradient bar reflects the loss of active compensation and

its degree, depending on the indicated genetic variation.



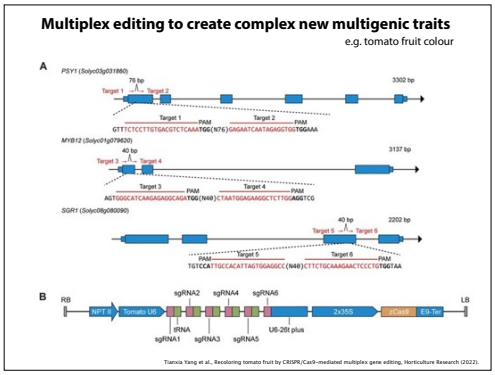
Clustering of CLE proteins from Brassicaceae and Solanaceae. Unlike the Brassicaceae (including Arabidopsis), the tomato genome contains a gene very closely related to CLV3, called CLE9, due to a genome duplication event [a]. Mutation of SlCLV3 results in a marked increase in fruit size. Mutation of SlCLE9 does not [b-e]. However loss of SlCLV3 function triggers a heavy increase in SlCLE9 expression by some kind of regulatory pertubation. This triggered overexpression of SLCLE9 partially compensates for the loss of CLV3 activity in the mutant. The heightened expression of SlCLE9 can be clearly seen in a CLV3 mutant background [c]. Further, double mutants of SlCLV3 and SlCLE9 show more extreme phenotypes than SlCLV3 mutants. *In vivo* editing of gene functions allows the unmasking of cryptic gene elements like SlCLE9, responsible for buffering in stem cell homeostasis in tomato.



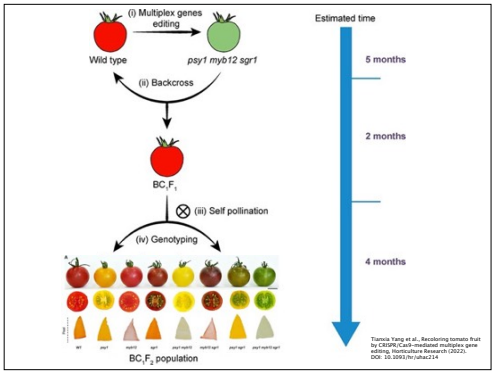
The compensating paralog SlCLE9 interacts with SlCLV3 in a sigmoidal dose-dependent epistasis relationship. Representative fruit images and locule-number quantification (mean ±1 SD) showing the effect of a Slcle9 mutant on locule number in WT plants and the Slclv3 promoter mutants. (https://doi.org/10.1126/science.adi5222)

Example: Multiplex gene editing

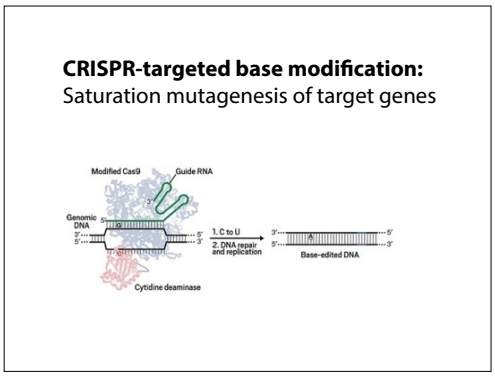
Simultaneous assorting of multiple gene functions



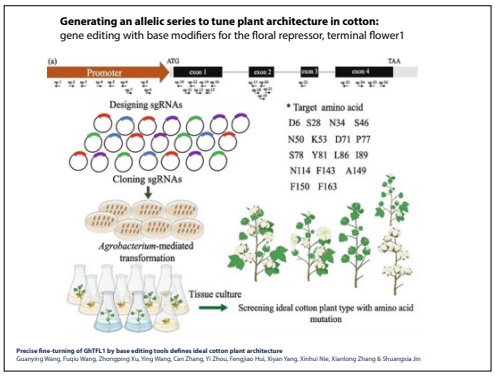
Tomato fruits display a wide range of colors, such as red, orange, pink, yellow, brown, green, purple, and white, which are determined by the levels and ratios of different pigments. The genes Phytoene Synthase 1 (PSY1), STAY GREEN 1 (SGR1) and R2R3-MYB transcription factor (MYB12) are associated with formation of different pigments. A range of psy1 sgr1 myb12 multiplex mutants were generated using CRISPR/Cas9 technology. Exon regions (blue) of the tomato fruit color-related genes were targeted by CRISPR/Cas9 cleavage and NHEJ repair. Letters in red represent the nucleotide sequence of the targeted regions, and letters in bold font indicate the protospacer-adjacent motif (PAM) sequence. The guide RNAs were delivered by a vector harbouring six sgRNAs in series. NPTII served as the resistance marker gene. All six sgRNAs and corresponding tRNA spacers were driven by the tomato U6 promoter and terminated by the U6-26t terminator. The Cas9 gene was driven by the 2×35S promoter. DOI: 10.1093/hr/uhac214



Schematic diagram shows the generation of tomato lines producing different colored fruits. The green-fruited psy1 myb12 sgr1 triple mutant was generated using the CRISPR/Cas9 system and backcrossed with the red-fruited WT cultivar 'Ailsa Craig'. Segregation of different mutant genes gave rise to a variety of fruit colours, stabilised as homozygous lines after self-fertilisation. DOI: 10.1093/hr/uhac214

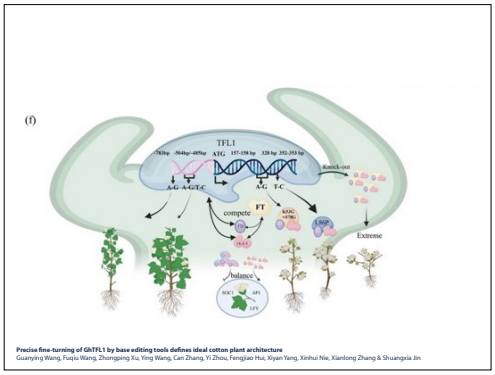


dCas9-CRISPR ribonucleoproteins can be used as RNA programmable tools for directing enzyme activities like base modifiers (e.g. cytidine deaminase) to specific targets in the genome. This allows targeted modification of particular coding regions, as the dCas9 complex is not capable of cleaving the DNA target - minimising deletions and insertions through non-homologous end-joining (NHEJ) reactions, yet capable of catalysing base changes for a wide range of mutant outcomes.



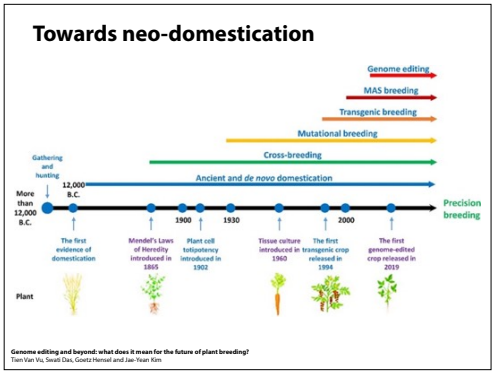
Artificially directed evolution of plant functional proteins through base editing. CRISPR-derived base editors have been constructed to engineer base changes in target genes. Cas proteins can be used to enzymes capable of converting bases adjacent to the targeting site(s). In this example, an adenine base editor (capable of catalysing A to G conversions) was targeted to the Terminal Flower 1 locus in *Gossypium hirsutum* (GhTFL1 in cotton). The TFL1 gene is highly conserved and plays an important role as an antiflorigen, regulating the switch to flowering in the shoot apical meristem. However, its complete inactivation using CRISPR/Cas9 knockout results in the appearance of extreme and agronomically unuseful traits. The non-coding and coding regions of GhTFL1 were targeted with 26 guide RNAs (shown

as black arrows under a schematic view of the gene). This allowed generation of a comprehensive allelic population of 300 independent lines. This allowed hidden pleiotropic roles for GhTFL1 to be revealed, and to create new alleles which confer moderate height, shortened fruiting branches, compact plants, and early-flowering.



Precise fine-turning of GhTFL1 by base editing tools defines ideal cotton plant architecture
Guanying Wang, Fuqiu Wang, Zhongping Xu, Ying Wang, Can Zhang, Yi Zhou, Fengjiao Hui, Xiyan Yang, Xinhui Nie, Xianlong Zhang & Shuangxia Jin

Model for antagonistic roles of TFL1 and Flowering Locus T (FT) in promoting branch or floral fate. Schematic model illustrating the role of GhTFL1. During the nutritional growth phase, GhTFL1 binds to Gh14-3-3 proteins, Flowering Locus D (GhFD) to repress the downstream flowering gene Apetala1 (GhAP1), thereby maintaining the nutritional growth of the plant. When the GhTFL1 gene is completely knocked out, nutritional growth is prematurely terminated, resulting in early flowering, reduced plant height, and other phenotypes. When a single base mutation occurs at some of the GhTFL1 loci, the function of GhTFL1 to repress downstream flowering genes is diminished, resulting in mutants with intermediate phenotypes.



Genome editing and beyond: what does it mean for the future of plant breeding?
Tien Van Vu, Swati Das, Goetz Hensel and Jae-Yean Kim

Plant breeding milestones. The start of domestication dates to around 12,000 BCE with the shift from hunting and gathering lifestyles to early agriculture. Discovery of Mendel's laws of genetics triggered new cross breeding and plant hybrid work. The discovery of totipotency of plant cells in the early 1900s was a precursor to *in vitro* tissue culture, introduced in 1960 with carrot. Plant tissue culture was a critical step for generating the first Agrobacterium-mediated transgenic plants in the early 1980's. In the meantime, mutational breeding using chemical or physical agents was introduced in the 1930s and played an important role in generating diverse genetic materials for crop breeding. Biochemical markers further enhanced crossbreeding in marker-assisted selection (MAS) breeding. Recently genome editing approaches have revolutionised directed plant breeding to a precision not seen before. Genome-edited soybeans with high oleic acid content were released in 2019, the first of many new crop varieties.

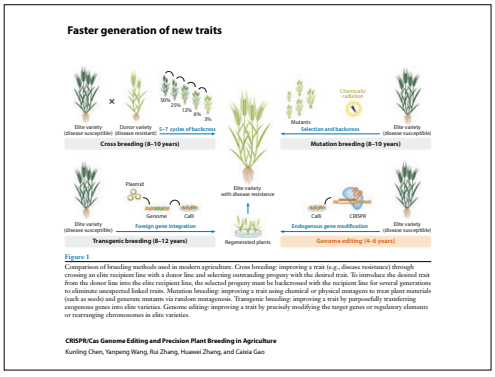
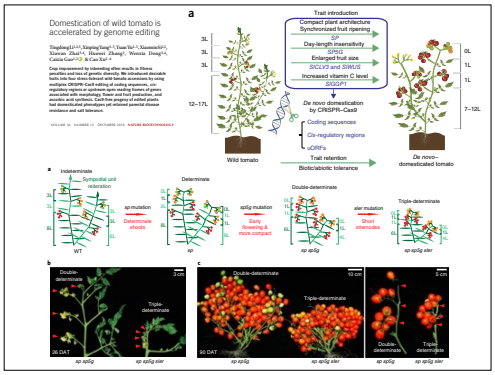


Figure 1
Comparison of breeding methods used in modern agriculture. Cross breeding: improving a trait (e.g., disease resistance) through crossing an elite recipient line with a donor line and selecting outstanding progeny with the desired trait. To introduce the desired trait from the donor line into the elite recipient line, the selected progeny must be backcrossed with the recipient line for several generations to eliminate unexpected linked traits. Mutation breeding: improving a trait using chemical or physical mutagens to treat plant materials (such as seeds) and generate mutants via random mutagenesis. Transgenic breeding: improving a trait by purposefully transferring exogenous genes into elite varieties. Genome editing: improving a trait by precisely modifying the target genes or regulatory elements or rearranging chromosomes in elite varieties.

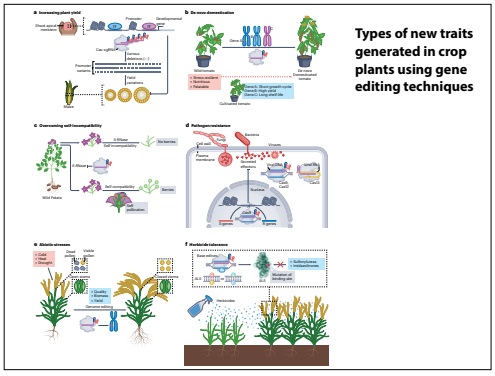
CRISPR/Cas Genome Editing and Precision Plant Breeding in Agriculture
Kunling Chen, Yanpeng Wang, Rui Zhang, Huawei Zhang, and Caixia Gao

Illustration of the advantages of speed and efficiency conferred by genome editing as a tool for plant breeding and design.

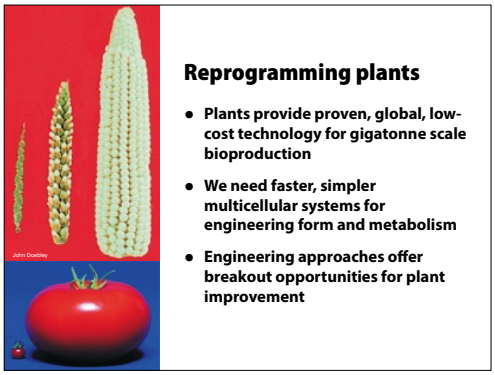


The stacking of CRISPR-Cas9 generated mutations has been used to recreate historic events in domestication of tomato. Regulatory elements of known genes have been specifically targeted to affect the genetic regulation of plant architecture, response to day length, fruit size, and fruit content - resulting in the de novo domestication of new tomato varieties. This demonstrates the feasibility of rapid domestication of other species - opening new opportunities for crop breeding, with faster speed and wider variety of traits.

Further, the same approaches can be used to create unknown traits, such as the example varieties shown with grape-like short floral internodes and hyper-clustered fruit, being explored for specialist types of urban farming.



There has been an explosion of academic and commercial work on CRISPR-Cas9 methods for targeted gene modification. The tools are being used to target genes involved in a wide variety of agronomic traits in a wide variety of plant species. The types of genetic systems being modified are represented here, and include those regulating crop yields, domesticating relatively unimproved species, altering breeding patterns, creating synthetic pathogen resistance, better tolerance of abiotic stresses and herbicide tolerance.



The editing or insertion of code to reprogram plant systems must be in the form of genetic (or epigenetic) information. The explosion of genetic analysis, genome sequencing and model plants provide tools and access to systems that (i) reach to lower level, molecular-scale networks, and (ii) upwards to modify cell and tissue scale growth. Over the coming decades, we are likely to see the emergence of comprehensive reprogramming of plant behaviour, and neo-domestication of new crops. Modern approaches to gene editing and reprogramming of the expression of master regulators allows rapid modification of plant architecture and other properties for crop improvement and possible domestication of new species.