

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.

1

The lecture images, handouts and references are available at:
<https://haseloff.plantsci.cam.ac.uk/education/index.html>

Lecture 1

1. Origins of modern crops
2. Selection and breeding of new crop varieties
3. Industrialisation of agriculture
4. Genetic modification (GM) for plant improvement
5. Genome editing
6. From science to engineering
7. Synthetic Biology

Lecture 2: Engineered DNA circuits.

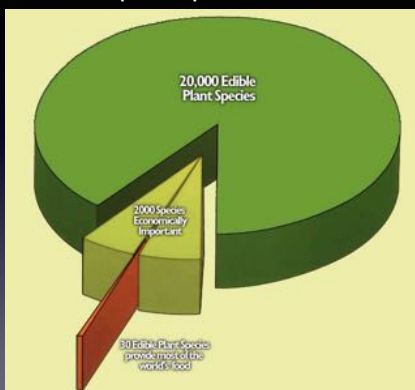
Lecture 3: Reprogramming of multicellular systems.

Additional resources: <https://haseloff.plantsci.cam.ac.uk> (Education)

2

Contents of this lecture

~400,000 plant species



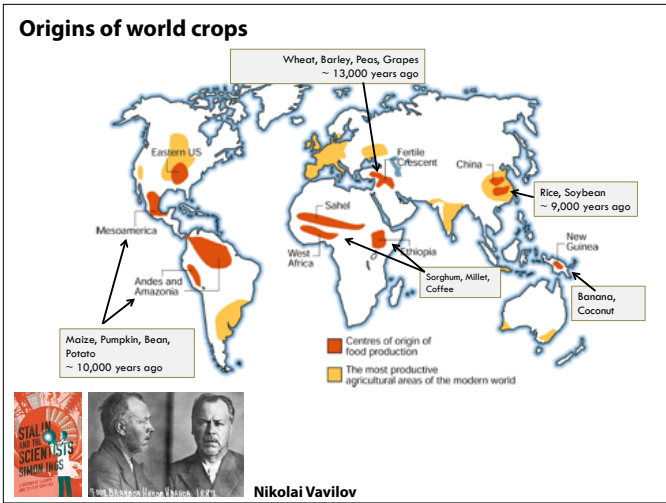
120 cultivated plant species

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

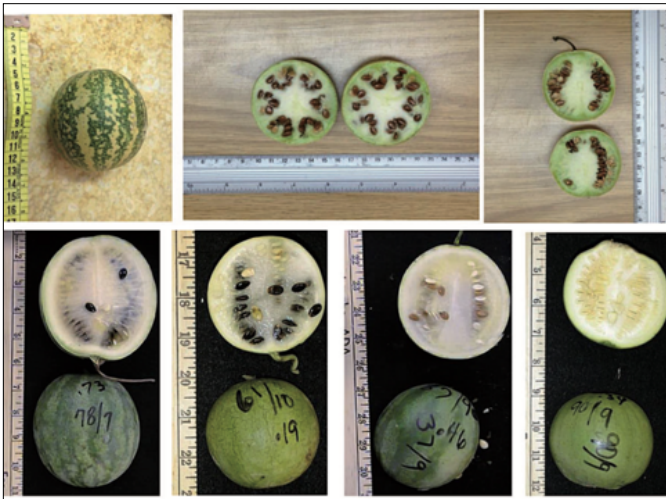
3

Crop plants sample a tiny fraction of total plant diversity. It is 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.

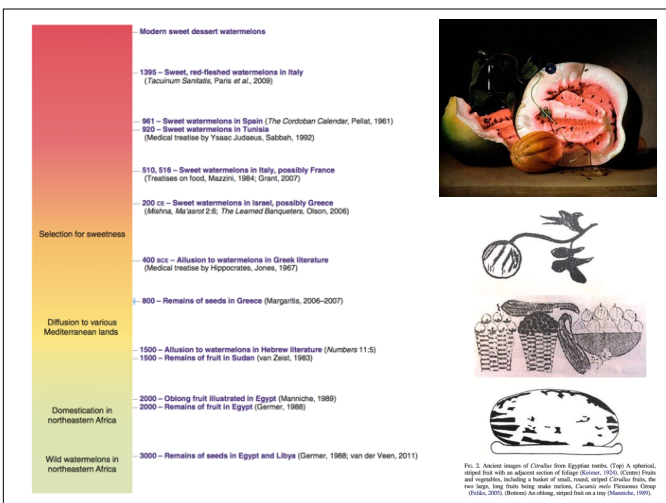
Nicolai Vavilov was a Russian biologist who first popularised the idea of geographical centres of diversity for the origin of modern crop species. These centres corresponded to areas of botanical diversity that coincided with the establishment of early human societies and plant domestication.



Ancient species are provided raw material for domestication of crop plants. Domestication has occurred over millennia, and often accompanied by substantial changes in phenotype. For example, the ancestor of the modern watermelon is believed to have originated in northern Africa. These ancestral plants possessed fruit that were pale, heavily seeded and bitter. The wild melons have a high water content but are bitter. However they were thought to have been originally used in prehistoric times as natural water carriers in northern Africa. The selection for sweeter tasting melons unintentionally produced pink flesh, as the genetic loci for colour and sweetness are closely positioned.



Images and the remains of watermelons have been found in 5000 year old Egyptian tombs. And there have been literary references to watermelons since that time. The first evidence of sweet watermelons occurs around 2000 years ago.





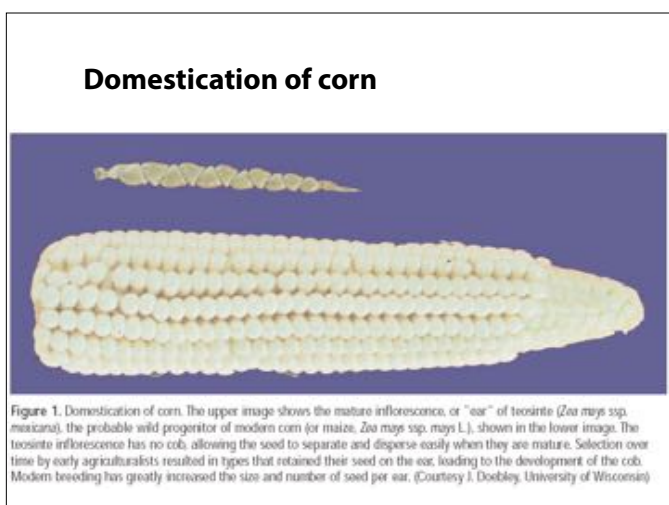
7

Watermelon phenotypes: ranging from the ancestral form (lower left), through to modern varieties.



8

A wide variety of modern cultivars are shown. Modern breeding has produced an expanded variety of different characters including fruit colour, size and seed content.



9

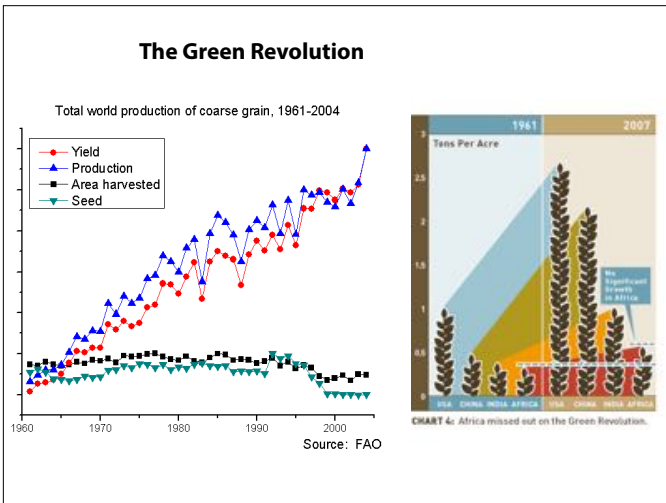
Early forms of maize strongly resemble teosinte, a plant endemic to Mesoamerica, and a subspecies of *Zea mays*. This likely progenitor has a strikingly distinct morphology, with smaller numbers of kernels arranged on a spike. It has been estimated that new varieties of maize been selected for over 9000 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

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: Doolley, University of Wisconsin)



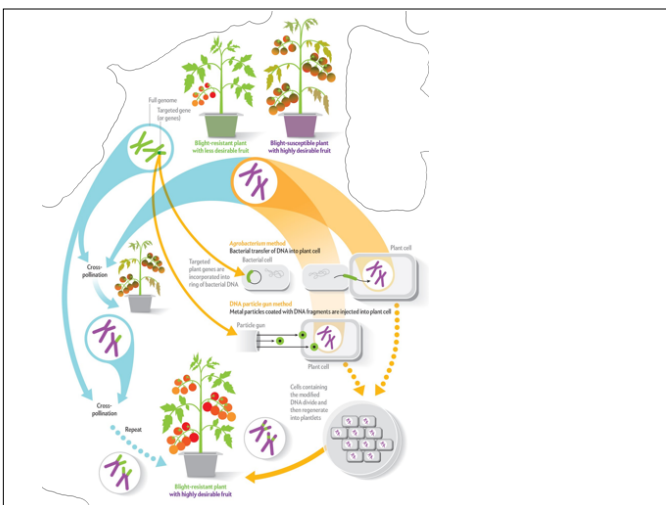
10

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.



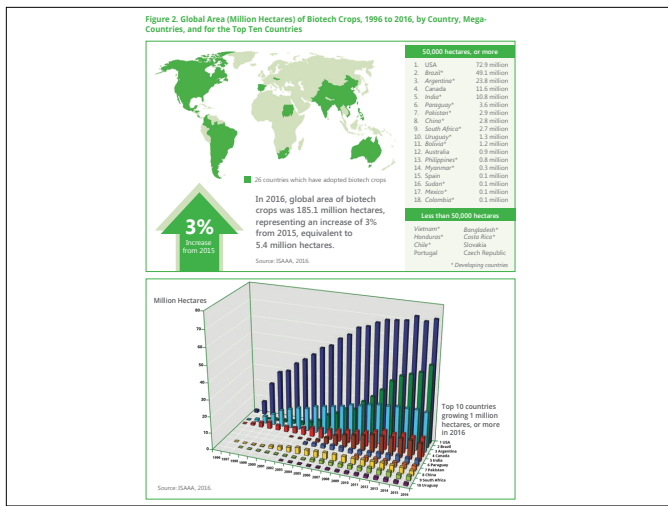
11

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.



12

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.



13

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.

Table 35. Global Area of Biotech Crops, 2015-2016: by Trait (Million Hectares)

Traits	2015	%	2016	%	+/-	%
Herbicide Tolerance	95.9	53	86.5	47	-9.3	-10
Stacked Traits	58.5	33	75.4	41	+16.9	+29
Insect Resistance	25.2	14	23.1	12	-2.1	-8
Virus Resistance/ Other	<1	<1	<1	<1	<1	<1
Total	179.7	100	185.1	100	+5.4	+3.0

Source: ISAAA, 2016

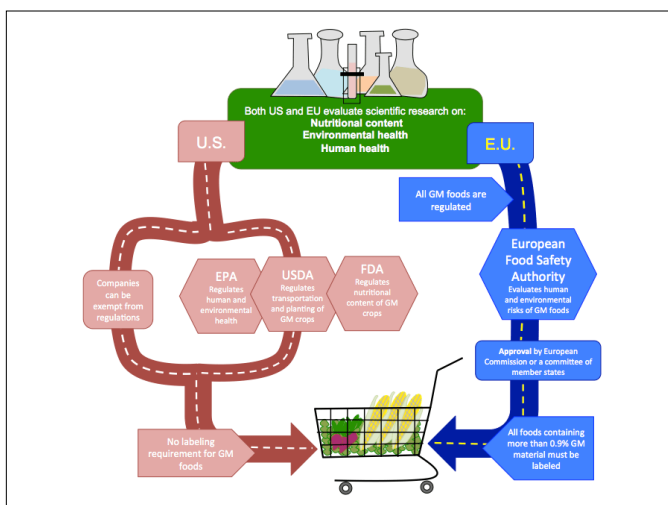
Table 29. Biotech Crop Area in the European Union, 2006-2016

Country	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
1 Spain	53,667	75,148	79,269	76,057	76,575	97,326	116,307	136,962	131,538	107,749	129,081
2 Portugal	1,250	4,263	4,851	5,094	4,868	7,724	9,278	8,171	8,542	8,017	7,069
3 Czechia	1,290	5,000	8,380	6,480	4,680	5,091	3,080	2,560	1,754	997	75
4 Romania	---	350	7,146	3,244	822	588	217	220	771	3	---
5 Slovakia	30	900	1,900	875	1,248	761	189	100	411	104	138
6 Germany	950	2,685	3,173	---	---	---	---	---	---	---	---
7 Poland	100	327	3,000	3,000	3,000	3,000	N/A	---	---	---	---
Total	57,287	88,673	107,719	94,750	91,193	114,490	129,071	148,013	143,016	116,870	136,363

Source: ISAAA, 2016

14

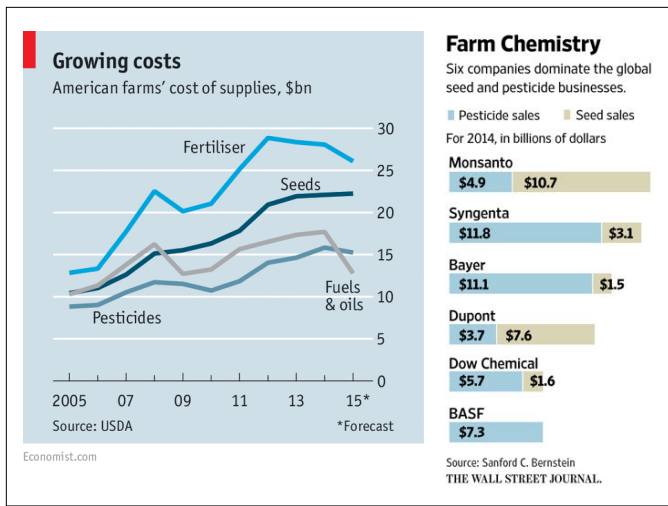
The first transgenic plants were created in the laboratory in the early 80s. By the mid-90s field trials of transgenic crops were underway. The first generation of traits included herbicide tolerance for weed control, and insect and virus pest resistance. In the subsequent 20 years there has been a rapid 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. Plus recent figures for the adoption of transgenic crops in Europe: there are relatively small areas of transgenic crops grown in Spain and Portugal - corresponding to transgenic maize. And very little grown elsewhere.



15

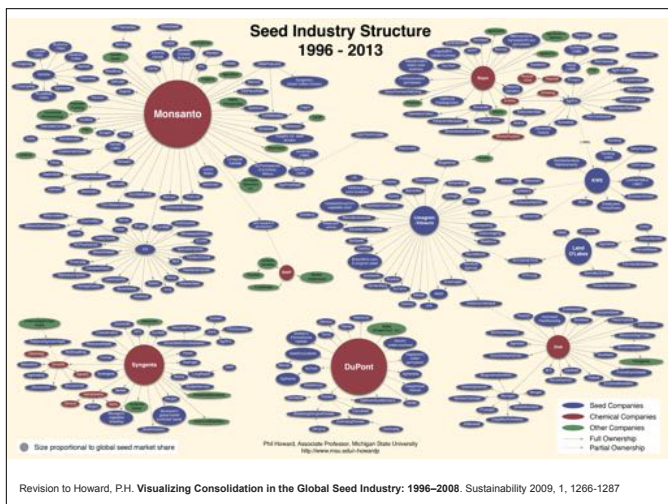
The US and Europe have adopted very different regulatory systems for GM foods. Food companies submit the same types of scientific data to U.S. and EU regulatory bodies for approval. Three separate agencies in the U.S. evaluate the potential risks of GM foods, while a centralised approval process is established 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: (i) Precautionary Principle (Europe)- GM crops are potentially dangerous and pose new risks and thus their use should be avoided until they are proven safe. (ii) 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.



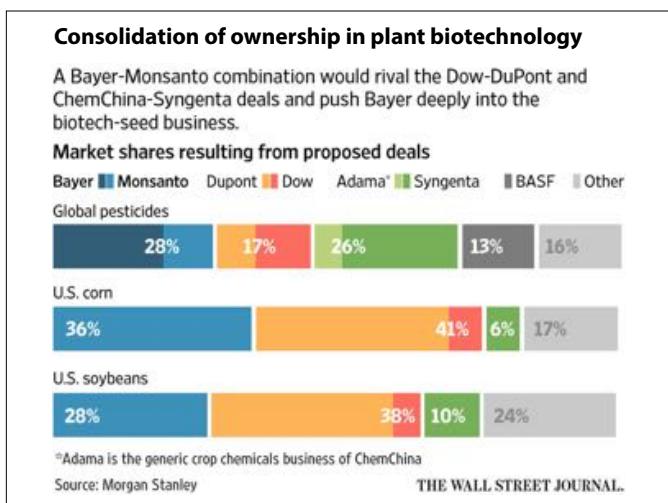
16

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.



17

Diagrammatic representation of the global seed industry. A few major agrochemical companies (shown in red) own, or have an interest in, clusters of the many seed companies. These agricultural combines are characterised by increasing vertical integration and consolidation.



18

Six major agrochemical companies are undergoing further mergers, and we may see three new companies owning 60% to 80% of key agricultural activities worldwide.

Disruptive technologies

Synthetic Biology: adoption of formal Engineering principles in Biology

19

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.

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)

20

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. But, first we will look at the emergence of engineering in different fields.

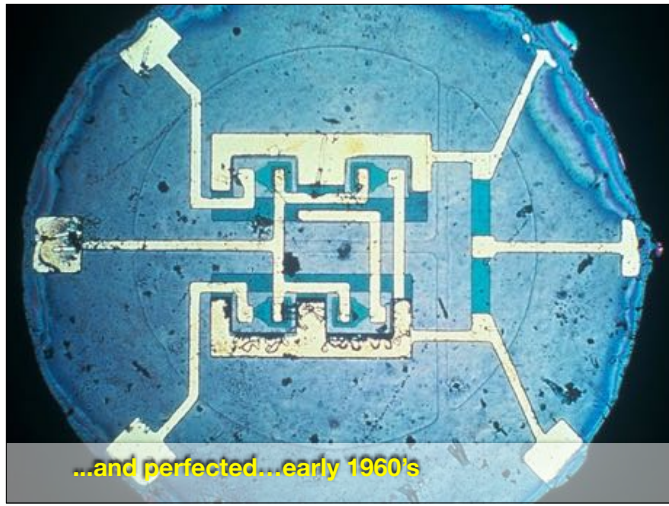
The Industrial Revolution was based on innovations in coal, iron, steam and mechanical engineering that took place in the mid-to-late 1700s. This led to inventions in the early 1800s, like the first modern steam engine found in Stephenson's Rocket. In mechanical engineering there was a lag phase between the periods of innovation and emergence of applications in manufacturing industries. The Industrial Revolution first took root in the United Kingdom. However major impacts on industrial output were first seen towards the mid-1800s. The application of new mechanical engineering principles in industry was accompanied by standardisation. For example, Whitworth proposed the first widely accepted standard screw threads for mechanical fasteners in 1830. Before this time, mechanical engineers needed to machine their own bespoke fasteners. The adoption of standards in mechanical engineering allow the use of interchangeable parts and facilitated the development of

21

Similar lags between innovation and impact were seen in other industries. The first microelectronic devices were crude and handmade.



1958 - First integrated circuit, Jack Kilby



22

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

1. Standardisation of parts

Standard mechanical and electrical interfaces were established for integrated devices by the early 1960's, and form the basis for today's microelectronics industry

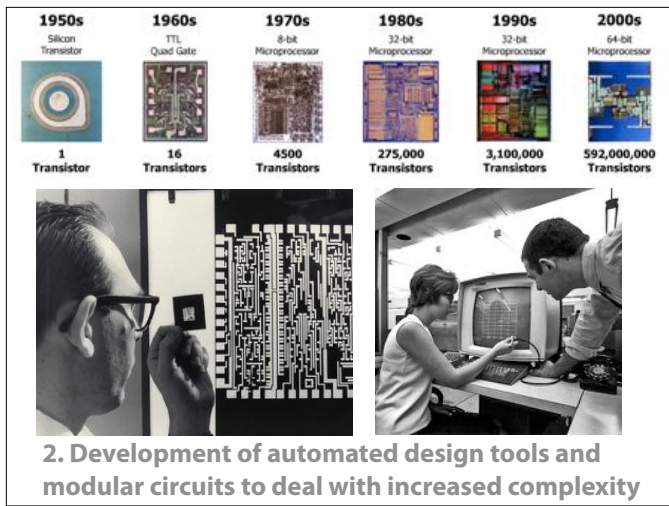
23

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.



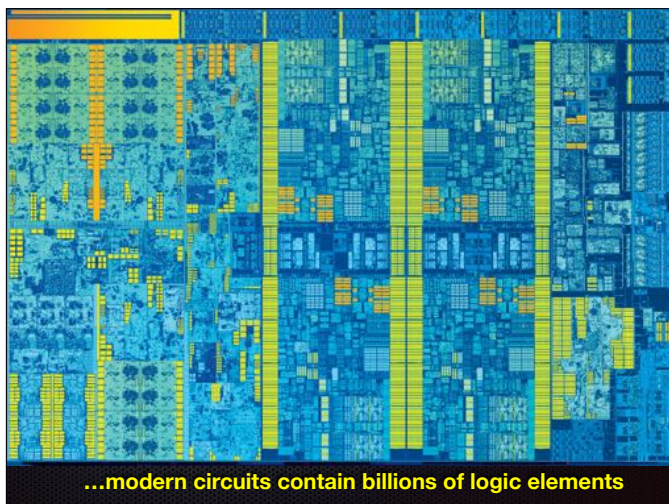
24

At first, these devices were designed by hand.



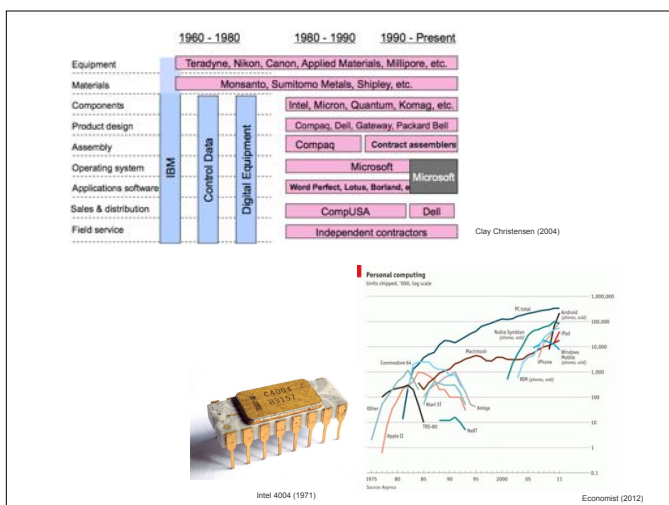
25

Increasing complexity saw the emergence of new automated design tools and reusable modular elements.



26

Modularisation and standardisation are the hallmarks of modern engineering. They allow management of highly complex systems.

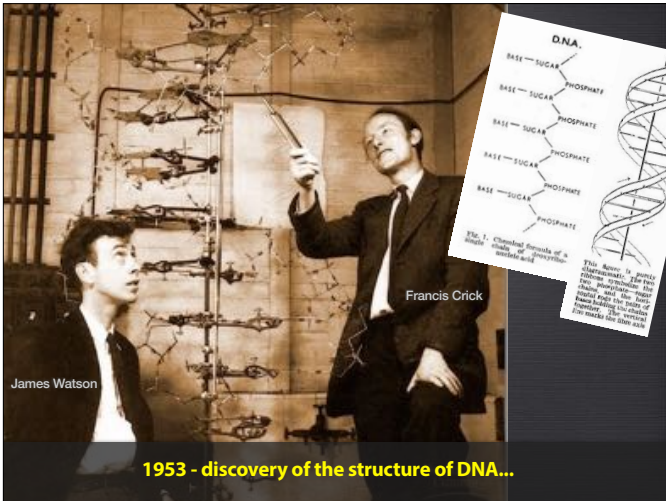


27

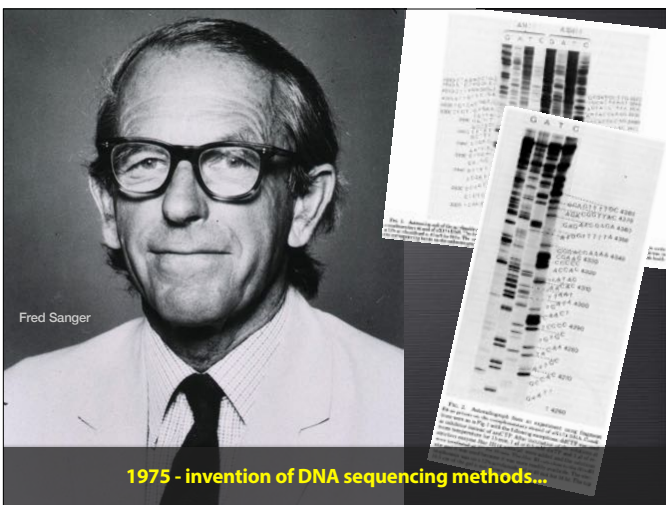
This level of consolidation has been seen in other industries. For example, the minicomputer industry was dominated by three companies (IBM, Control Data and DEC) through the 1960s. However the invention of the microprocessor in the early 70s, and the emergence of low-cost microcomputers cause disruption 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?

Meanwhile...

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...



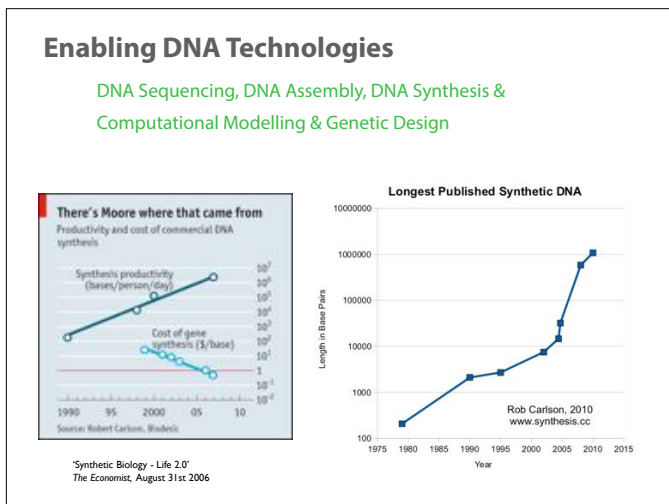
... To the development of DNA sequencing methods - at the kilobase-scale with Sanger sequencing in 1975





31

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



32

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.

Construction of Biologically Functional Bacterial Plasmids *In Vitro*

(R factor/restriction enzyme/transformation/endonuclease/antibiotic resistance)

STANLEY N. COHEN*, ANNIE C. Y. CHANG*, HERBERT W. BOYER†, AND ROBERT B. HELLING†

* Department of Medicine, Stanford University School of Medicine, Stanford, California 94305; and † Department of Microbiology, University of California at San Francisco, San Francisco, Calif. 94122

Communicated by Norman Davidson, July 18, 1973

ABSTRACT The construction of new plasmid DNA species by *in vitro* joining of restriction endonuclease-generated fragments of separate plasmids is described. Newly constructed plasmids that are inserted into *Escherichia coli* by transformation are shown to be biologically functional replicons that possess genetic properties and nucleotide base sequences from both of the parent DNA molecules. Functional plasmids can be obtained by reassociation of endonuclease-generated fragments of larger replicons, as well as by joining of plasmid DNA molecules of entirely different origins.

1973 - first molecular cloning experiments...

33

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

34

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



Smithsonian museum exhibit: Stanley Cohen's laboratory bench

35

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>

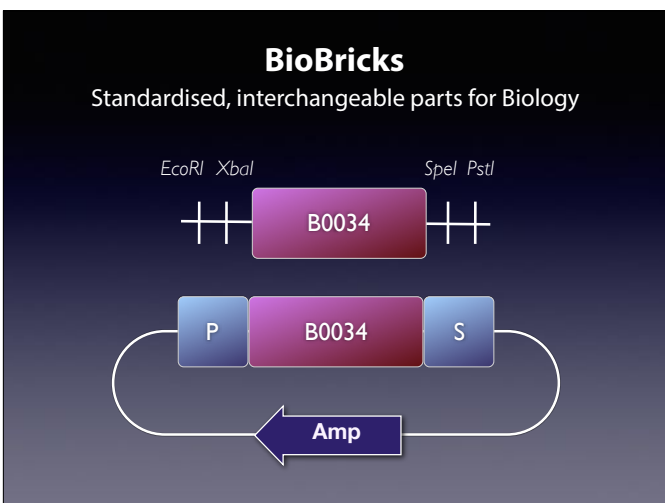


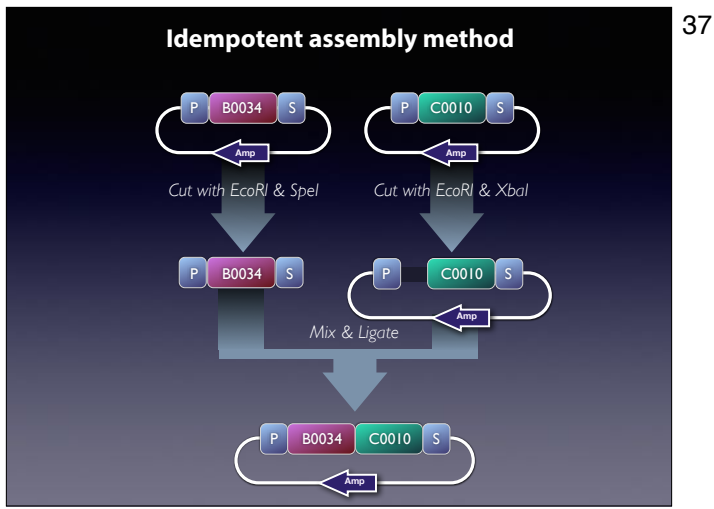
Tom Knight, MIT

Invention of standardised parts for biology...

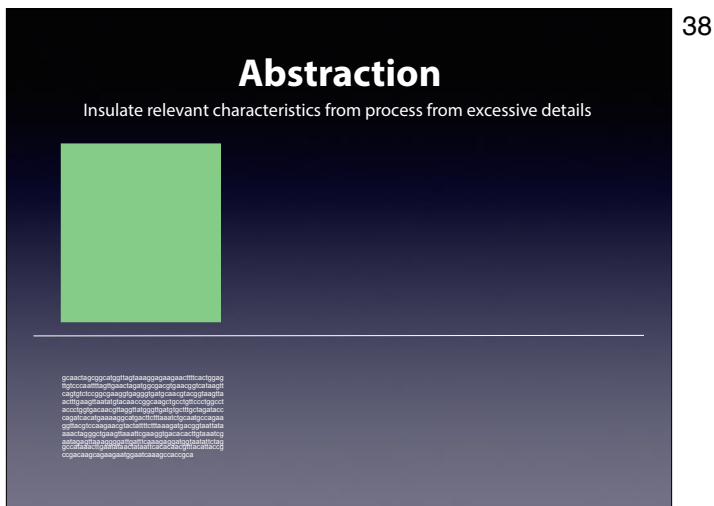
36

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

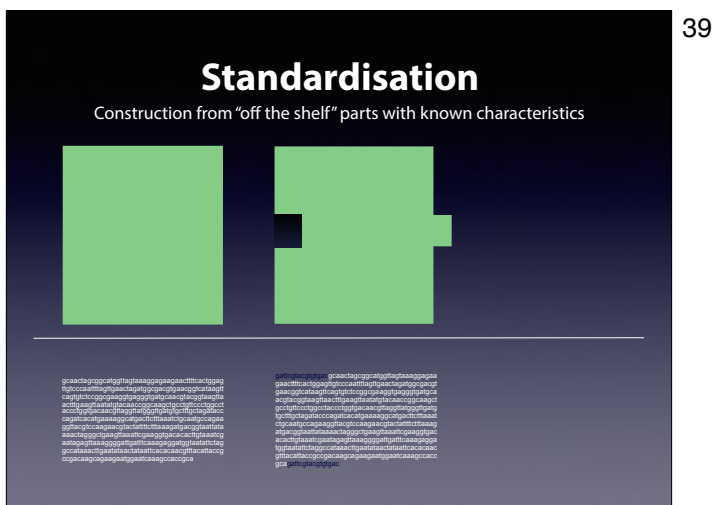




...the combination of any two parts would recreate the format of a standard part. (An object's properties remains unchanged during an idempotent operation). Note the arrangement of prefix (P) and suffix (S) elements in this diagram, as two fragments are ligated.



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.

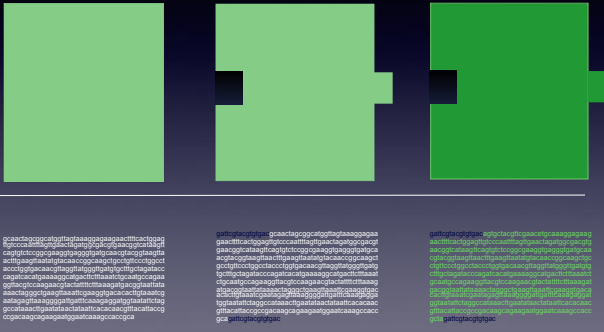


Standardised parts come with an implied means of assembly. They can "plug" together in a manner similar to Lego parts.

40

Decoupling

Insulate design process from fabrication details



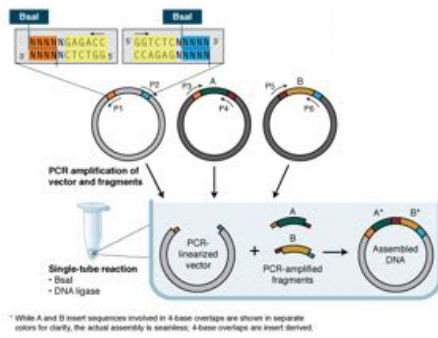
The process of improvement of DNA parts can be separated from the design process. In this case, a new part with a modified coding sequence for a brighter green fluorescent protein can be used interchangeably by a genetic circuit designer. The design process is decoupled from the fabrication process.

41

A revolution in DNA assembly

Type IIS restriction enzyme assembly techniques (e.g. Golden Gate, MoClo, Golden Braid, Loop Assembly):

- (i) Multiplex assembly, with simultaneous precise joining of multiple sequences;
- (ii) Highly efficient ligation, as side products are rescued and reactions pushed to completion;
- (iii) No need to purify DNA fragments for ligation, as these are generated during the reaction;
- (iv) Single tube reaction, easy to automate;
- (v) Standardisation of DNA parts and vector composition.



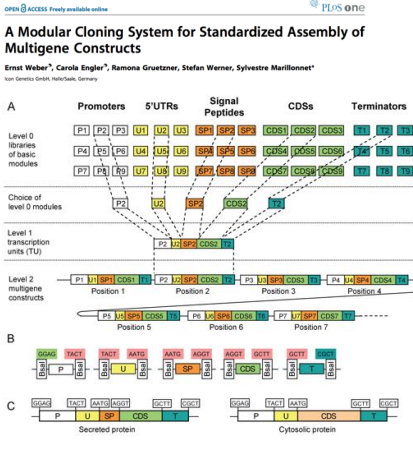
This approach has become increasingly sophisticated, now in the form of type IIS assembly techniques. These rely on restriction enzymes with cleavage sites that are offset from their recognition sequence. There is no need to isolate DNA fragments. Intact plasmid DNAs can be mixed, and cleavage and ligation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

42

**Type IIS
DNA assembly
protocols:**

**Golden Gate
MoClo
ENSA
Golden Braid
Loop assembly:**

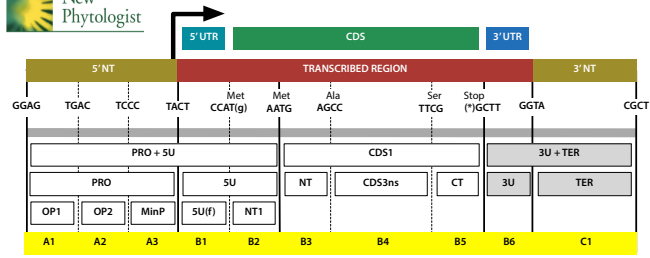
**adopted by the
plant research
community**



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.

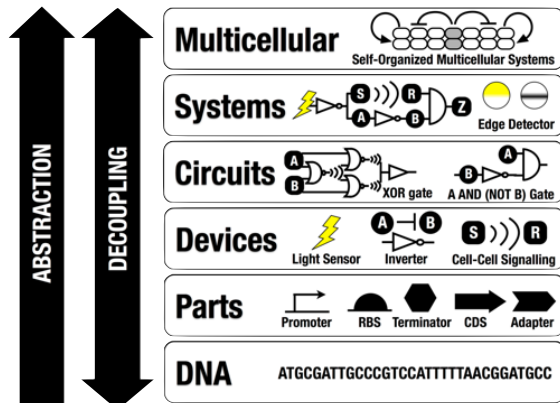


NPH-L-2015-19556.R1 Standards for Plant Synthetic Biology: A Common Syntax for Exchange of DNA Parts
by Patroy, Nicola; Orzaez, Diego; Harilal, Sylvester; Warscha, Herbert; Matthouan, Colette; Youles, Mark; Raitskin, Oleg; Leveau, Aymeric; Ferrer-Martinez, Gemma; Rogers, Christian; Smith, Alison; Hibberd, Julian; Webb, Alex; Locke, James; Schornack, Sebastian; Ajlona, Jim; Boultonbe, David; Zipfel, Cyril; Kamoun, Sophie; James, Jonathan; Kahn, Hamam; 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; Graneli, Antonio; Tissier, Alain; Shih, Patrick; Bruteau, Thomas; Buick, Paul; Bischof, Heiko; Fraser, Paul; Aharoni, Asaph; Raines, Christine; South, Paul; Ané, Jean-Michel; Hamberger, Björn; Langdale, Jane; Stougaard, Jens; Bouwmeester, Harro; Udvardi, Michael; Murray, Jim; Ntoukakis, Vardis; Schäfer, Patrick; Benby, Katherine; Edwards, Keith; Osbourn, Anne; Haseloff, Jim; Haseloff, Jim

43

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

44

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.

Lecture 1

1. Origins of modern crops
2. Selection and breeding of new crop varieties
3. Industrialisation of agriculture
4. Genetic modification (GM) for plant improvement
5. Consolidation and vertical integration
6. From science to engineering
7. Synthetic Biology

Lecture 2: Engineered DNA circuits.

Lecture 3: Reprogramming of multicellular systems.

Additional resources: <https://haseloff.plantsci.cam.ac.uk> (Education)

45

Contents of this lecture

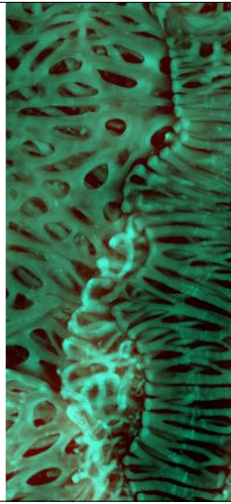
Genomics, Epigenetics & Synthetic Biology

Synthetic Biology and Plant Biotechnology: Lecture 2

Engineered DNA circuits

Jim Haseloff

<https://haseloff.plantsci.cam.ac.uk>



1

Lecture 2

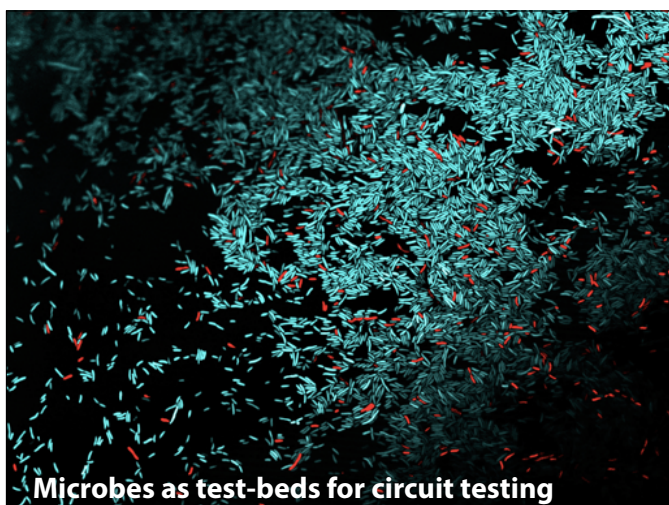
1. Microbial test systems
2. Feedback-mediated regulation
3. Repressilator
4. Circuit optimisation
5. *In vitro* systems for rapid testing
6. Single cell analysis
7. Computational modelling
8. Megabase-scale genome refactoring
9. Engineering of cellular metabolism

Lecture 3: Reprogramming of multicellular systems.

Additional resources: <https://haseloff.plantsci.cam.ac.uk> (Education)

2

Genetic circuits can be divided into two classes. The first consists of circuits which act cell autonomously. These include regulatory circuits that might act as sensors, or metabolic circuits that affect local cell properties. The second includes circuits that allow communication between cells and might have self organising properties. This lecture will describe the challenges and current state of the art for design and assembly of DNA circuits.

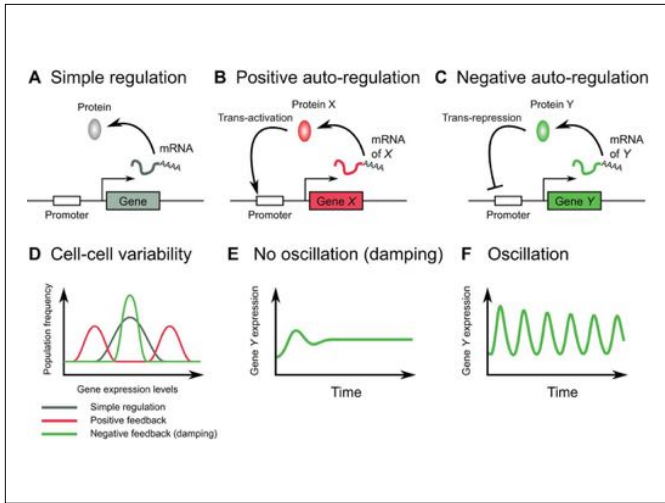


Microbes as test-beds for circuit testing

3

The first experiments in synthetic biology have focused on the use of microbes and *in vitro* expression systems - a simple and facile testbeds for testing new circuits.

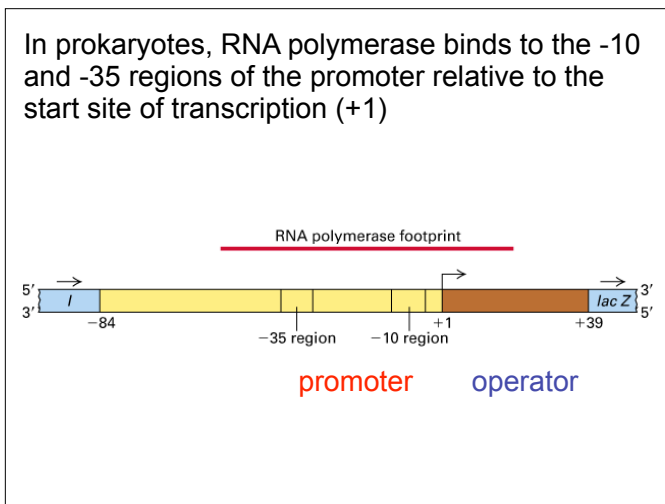
4



Regulatory elements can be programmed into DNA to produce circuits capable of a wide range of behaviours. These include simple gene expression or autoregulation to induce or repress gene activity. These basic regulatory activities can be combined to control cell to cell variability, induce gene expression and create oscillatory behaviour.

5

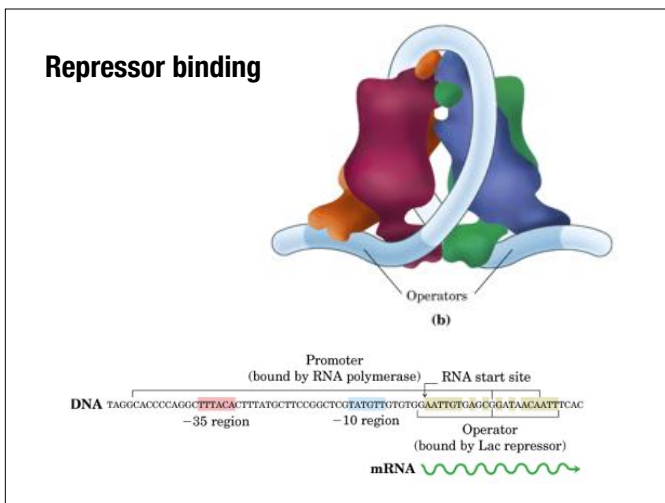
In prokaryotes, RNA polymerase binds to the -10 and -35 regions of the promoter relative to the start site of transcription (+1)



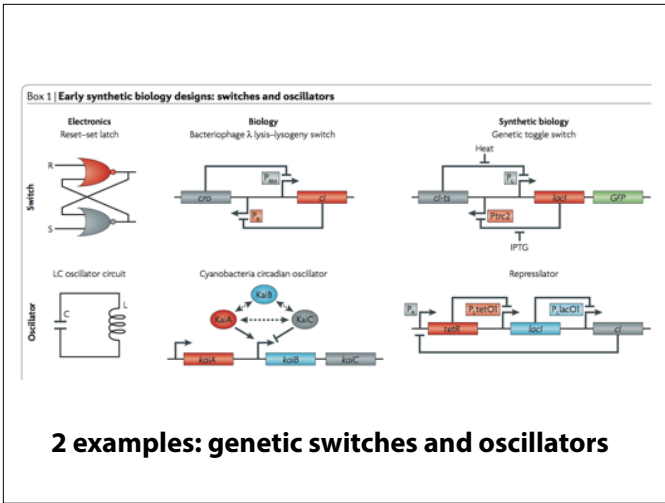
The building blocks for the circuits come from components of normal microbial gene expression. Transcriptional control plays a major part in regulation of bacterial gene expression. Recruitment of RNA polymerase is regulated by sequences adjacent to the start site for transcription. Regulatory proteins can either inhibit binding or recruit the RNA polymerase to the promoter sequence. These repressor or activator proteins bind to what are termed operator sequences. The addition of an operator to a promoter will confer ectopic regulation to that gene.

6

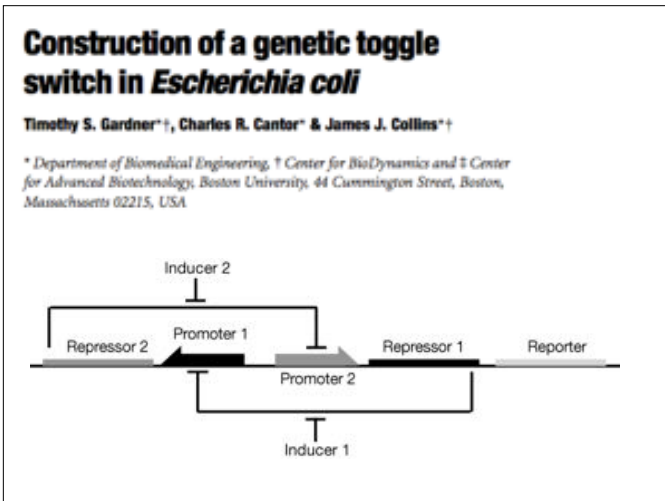
Repressor binding



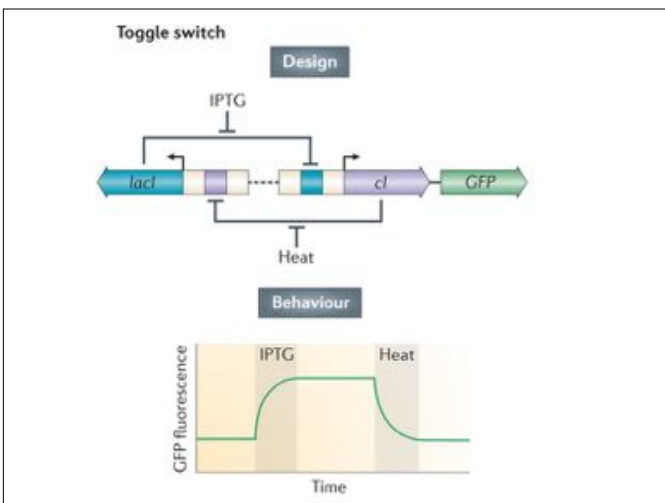
Lac repressor proteins binds as a dimer to adjacent inverted sequences in an operator. The resulting Lac dimer is capable of binding to another dimer. The presence of a second operator in a gene will allow the formation of a DNA loop structure, and tight repression of gene expression.



These types of regulatory regions can be used to create synthetic circuits composed of DNA. We will look at two examples. The first shows switch-like behaviour, and second shows oscillatory behaviour. These synthetic circuits have counterparts both in electronic logic circuits and in natural biological circuits.



A simple switch can be built from two repressor genes, which are each configured to repress the expression of the other. In this circuit, small molecule inducers are used to regulate the behaviour of each repressor. For example, the presence of inducer 1 will allow expression from promoter 1 therefore accumulation of the repressor 2 protein - and therefore production of repressor 1 will be stably inhibited - even in the absence of the inducer. If inducer 2 is present, repressor 1 will be produced, and this will maintain a repressed state for repressor 2. In other words there are two metastable states, and gene expression can toggle between either.



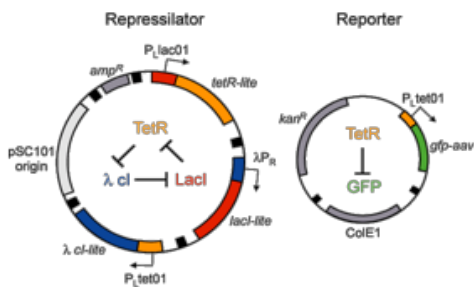
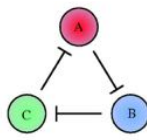
Schematic diagram of a genetic toggle switch built with genes for the lambda repressor (cl) and Lac repressor (lacI). Inducibility is provided by IPTG for the Lac repressor, and use of a temperature sensitive variant of cl. The state of the genetic switch is indicated by expression of a fluorescent protein linked to expression of the Lambda repressor. This is either on or off. Once switched, the circuit retains its existing state until exposed to the opposite inducer. Hence it behaves like a conventional light switch.

The Repressilator

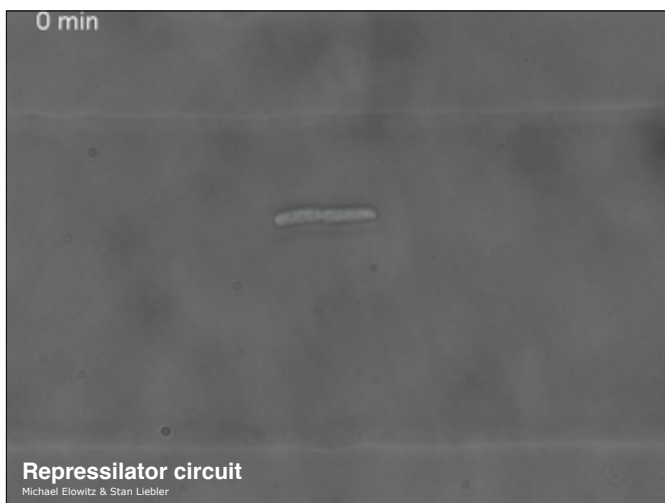
- Cellular clocks oscillate with defined periods
 - Circadian clocks oscillate with 24-hour period
- Elowitz and Leibler set out to build oscillator with components not found in cellular clocks
- Used three transcription factors in mutual-repression network
 - Lacl
 - TetR
 - cl from lambda phage
- Readout: GFP controlled by Tet repressor

Elowitz and Leibler decided to try to build a synthetic oscillator without using any of the known components of existing clocks. They chose three well-characterized transcription factors: the lac repressor, which prevents transcription from the lac operon in *E. coli* bacteria, the Tet repressor that prevents transcription from the Tet operon, and the λ repressor that prevents transcription from the Tet operon, and the λ repressor acts in lambda phage to regulate the choice between the lytic and lysogenic phases. These three genes were organized in a mutual-repression network. As a readout for the oscillator, a TetR regulated GFP (green fluorescent protein) was used.

Repressilator design



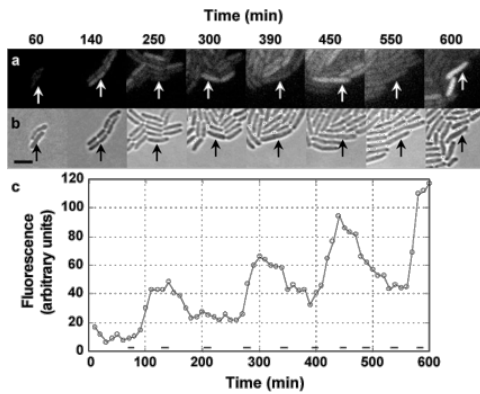
Each of the transcription factors acts to repress transcription from another transcription-factor gene: TetR represses lambda cl, lambda cl represses Lacl, and Lacl represses TetR. As the concentration of TetR increases, it shuts down expression of lambda cl, which reduces the concentration of the lambda cl repressor in the cell, which allows Lacl expression to increase. However, the resulting increase in Lacl repressor acts to shut down TetR expression, which in turn causes an increase in lambda cl, which represses Lacl, etc. Thus, the system oscillates. To visualize the oscillatory behavior of the network, the investigators used GFP controlled by TetR. This construct was carried on a separate plasmid. To reduce the period of oscillation of the network, the repressor proteins were genetically destabilised.



Visualisation of the repressilator circuit in action in individual cells. This first version of the repressilator showed poor synchronisation as the period and amplitude of the oscillations were susceptible to variation due to noise during cell proliferation.

Activity of repressilator

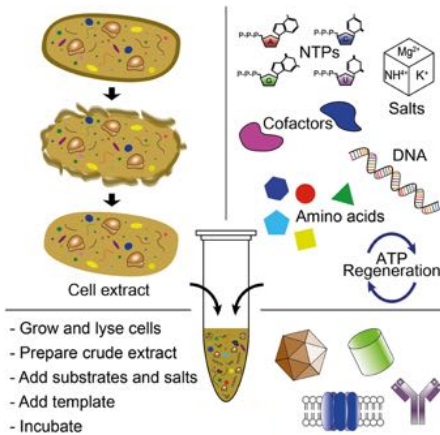
13



Oscillatory behavior of the repressilator. In the upper panel, a single *E. coli* bacterial cell is followed as it divides to form a microcolony. In the lower portion of this panel are bright-field images showing the increase in the number of bacteria. The upper part of the panel shows the GFP fluorescence readout of the repressilator, which has a period of approximately 150 minutes. Thus, at 60 minutes there is a low point of activity, at 140 minutes a peak of fluorescence, at 250 minutes another trough, at 300 minutes another peak, etc. In the graph in the lower panel, the level of fluorescence is plotted against time. The small bars at the bottom of the graph show the times when the bacteria were dividing. The period of the oscillations of the repressilator do not correspond to the bacteria's division cycle. Thus, the increases and decreases of the proteins involved in repressilator function continue independently of the division cycle of the bacteria.

The parameters associated with gene production rates, protein properties and DNA affinities can be used to create simple continuous models for behaviour of the repressilator circuit. Computer models used to predict the behaviour of the circuit,

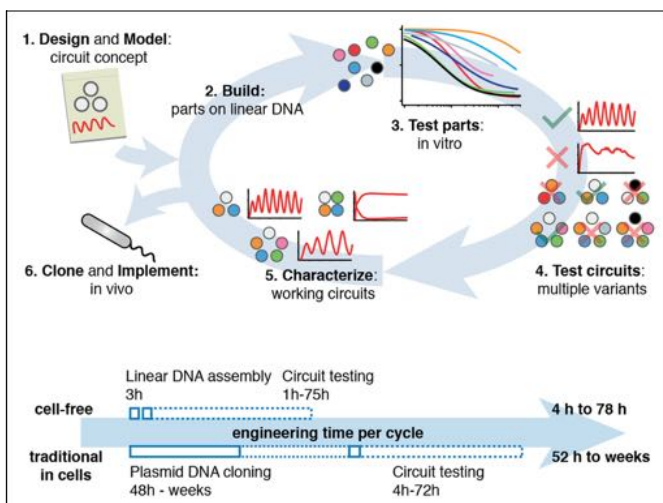
14

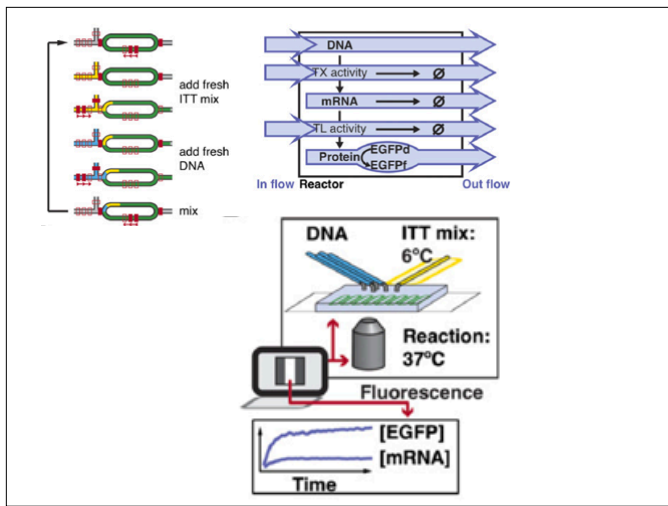


Cell-free extracts of *E. coli* have been used to characterise circuits in vitro. The cell-free extracts contain transcription and translation components from the bacterium, to which are added nucleotides, amino acids salts and an energy production system. The extracts contain inverted vesicles derived from the plasma membrane. These act to sustain energy production in the extract by oxidative phosphorylation. The extracts can then be programmed by added DNA.

15

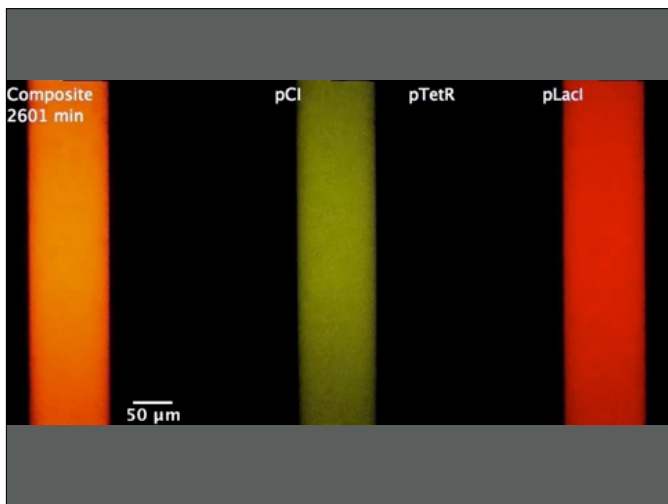
...and can be used to accelerate the testing of circuit variants at small-scale. *In vitro* cell extracts have been used to test and improve variants of the repressilator circuit.





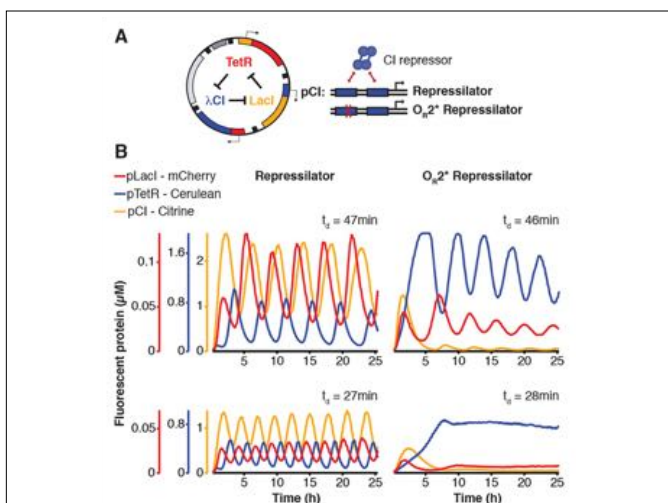
16

In work from the Maerkl Lab, has exploited use of microfluidic containers to feed reactants into a central chamber where repressilator circuit was active. Valves and pumps used to mimic the process of DNA replication and cytoplasmic dilution through cell division.



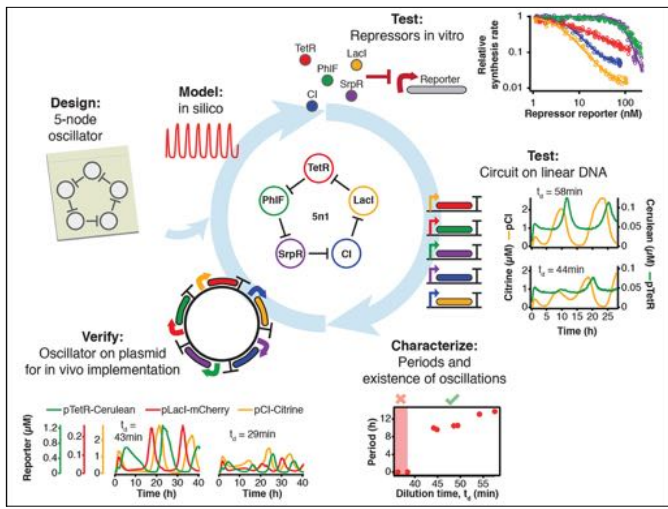
17

An example of improved repressilator activities, here shown with multispectral fluorescent protein outputs.



18

It was then relatively simple to adjust the reaction conditions and DNA templates to systematically explore the behaviour of many circuit variants and conditions. Here showing the behaviour of the repressilator when different strength operators were used. The biochemical observations and parameter measurements were paired with the use of mathematical models for the circuit.



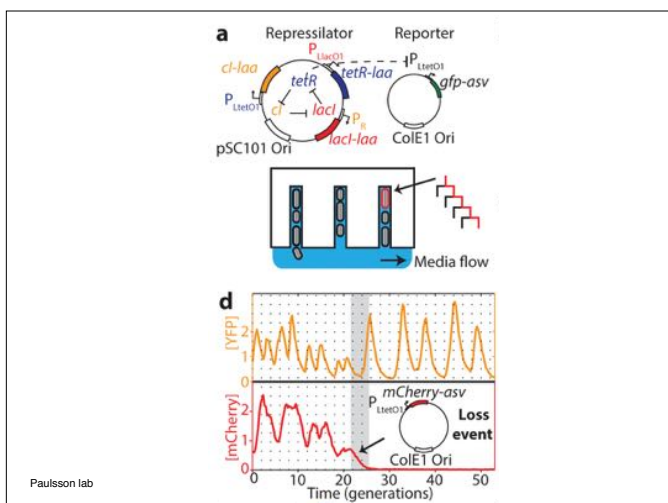
19

This Design-Build-Test approach using *in vitro* gene expression was used to test larger scale circuits (5 repressors). An odd number of elements in the gene circuit gives rise to oscillations.



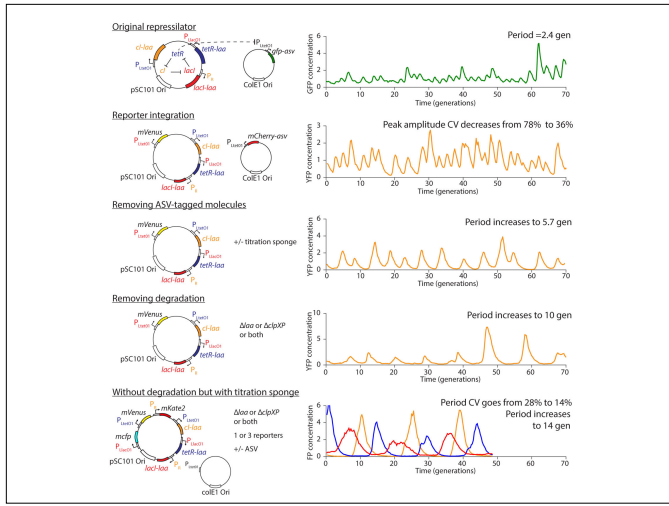
20

The circuits which showed improved robustness *in vitro* could be implemented in cells also. Here, the Maerkl lab used a "mother machine" - where cells are trapped in narrow chambers, and cell divisions and lineages can be tracked precisely. Dividing cells largely remain in synchrony.



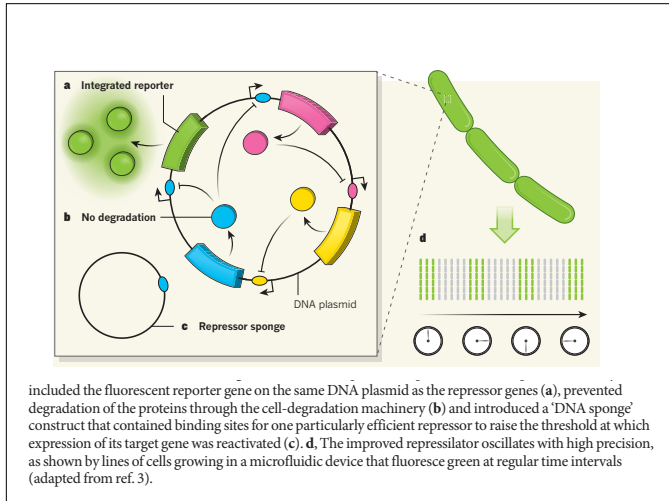
21

The Paulsson lab further used "mother machine" based measurements to improve the basic circuit. They discovered that various types of molecular competition and cross-talk in the circuit were sources of noise. For example, the reporter GFP gene was carried on a high copy number plasmid - separate from the repressilator circuit. Loss of the reporter plasmid was observed to cause more regular, higher amplitude cycles and longer period. This was shown to be due to competition for protein degradation machinery - as all repressors and GFP reporter shared 3-amino acid degradation tags.

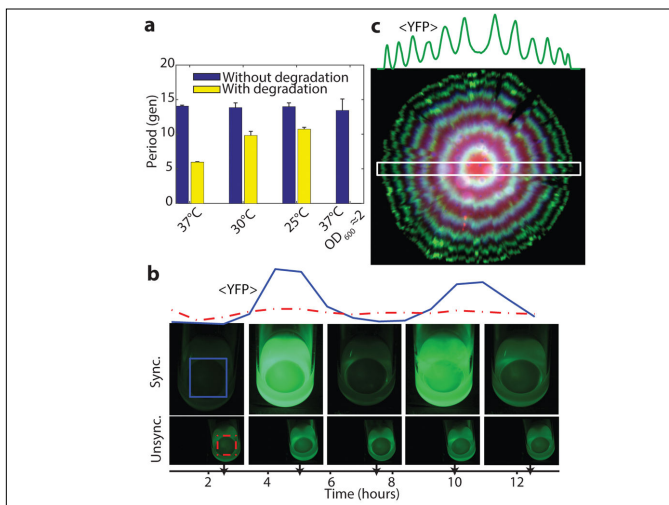


22

More robust cycles were seen when this and similar other sources of noise were removed. (i) Integration of the reporter gene on the low-copy repressilator plasmid. (ii) Removal of all destabilised (ASV C-terminus tagged) proteins). (iii) Use of a mutant E. coli strain with defective protein turnover. (iv) Addition of a plasmid borne "sponge" for binding excess TetR protein.



23



24

The combination of these improvements gave rise to cell lines that maintained long period oscillations that remained in phase over many cell generations. Synchronised liquid cultures showed bulk oscillatory behaviour. Colonies grown on solid media showed ring-like features as the cells grew outward from the point of inoculation.

Design and assembly of large scale genetic circuits

- Problem: to understand design principles of biological networks
- Approach: Design and construct synthetic network
- Knowledge gained from the design of synthetic networks should help understand real networks
- Could be used to engineer new cellular behaviours in plants

A key problem facing the field of systems biology is in understanding how biological networks are put together, or, in systems engineering terms, understanding their “design principles.” One approach is to design and construct synthetic networks and then see how the synthetic networks function when introduced into cells. Once introduced into an organism, a synthetic network can be subjected to various perturbations to see how the network responds. In addition to helping us understand real cellular networks, synthetic networks have the potential to be used to engineer new cellular behaviors.

In addition to the assembly of genetic systems from standardised parts, we are also seeing the re-factoring, synthesis and transplantation of entire genomes.

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

The entire *Mycoplasma mycoides* genome was synthesised in the form of small oligonucleotides. These were stitched together to form one kilobase-sized cassettes, which were in turn assembled into first 10kb, and then 100 kb fragments. These were finally assembled into the megabase circular genome of the bacterium. Small fragments were assembled *in vitro*, the larger fragments were assembled by homologous recombination *in vivo*, using yeast as a host.

Chemical synthesis of a Bacterial Genome: *Mycoplasma mycoides*, JCVI 2010

Assembly of *M. mycoides* genome

1. Overlapping oligonucleotides (including yeast vector, *lacZ*) recombined to make 1080 bp cassettes (orange arrows).
2. In sets of 10, the cassettes recombined to produce 109 ~10kb assemblies (blue arrows)
3. In sets of 10, the ~10kb assemblies recombined to produce 11 ~100kb assemblies (green arrows)
4. These 11 assemblies were recombined to the final genome, sMmYcP235 synthetic genome to create JCVI-syn1.0 cell line

Large DNA fragment assembly via homologous recombination in yeast

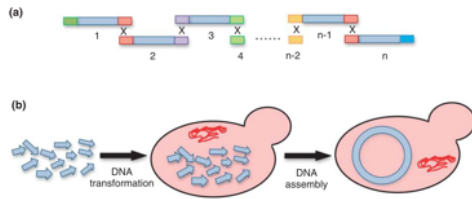
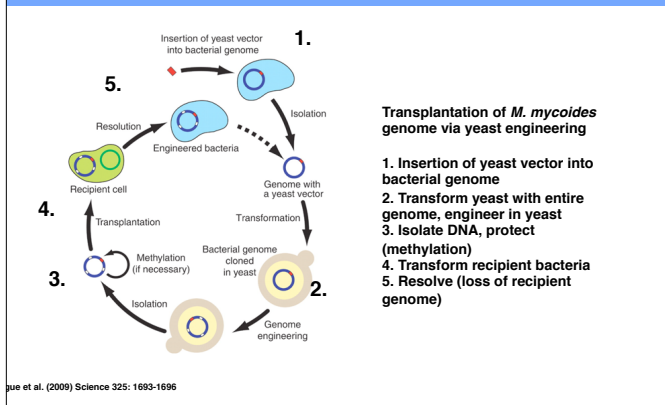


FIGURE 2 | Synthesis of large DNA molecules in yeast. (a) Yeast homologous recombination mechanism. DNA fragments sharing an overlap region at 3'- and 5'-ends with the neighboring DNA fragments can be assembled into a single larger DNA molecule. (b) Construction of a synthetic *M. genitalium* genome. Twenty-five different overlapping DNA segments (blue arrows, 17–35 kb each) composing the genome were co-transformed into yeast followed by assembly of the entire genome in a single step.

The task of assembling a complete bacterial genome was made possible by a number of technical innovations. First, an efficient technique for the multiplex assembly of scar free DNA fragments was developed. This was called Gibson assembly, after the inventor. Second, large DNA fragments could be assembled in an ordered fashion using homologous recombination in yeast. This technique is capable of producing chromosome scale synthetic DNAs.

Creating Bacterial Strains from Genomes that have been cloned and engineered in yeast



Third, the judicious inclusion of sequences that allowed replication and maintenance of foreign DNAs in yeast - allowed the propagation, manipulation isolation of chromosome-scale synthetic DNAs. The synthetic bacterial chromosomes could then be isolated from these yeast strains.

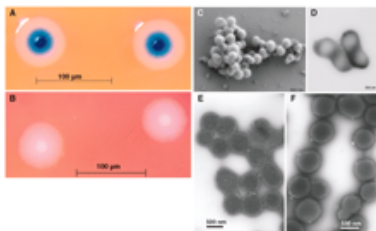
Transplantation of a synthetic bacterial genome: *Mycoplasma mycoides*, JCVI 2010

Genome transplantation

DNA from the final assembly in yeast sMmYCp235 synthetic genome was transplanted into a *M. capricolum* cell to ultimately produce JCVI-syn1.0

A&B. WT *M. mycoides* colonies are white, JCVI-syn1.0 are blue (*lacZ*beta galactosidase + Xgal). "Fried egg" morphologies characteristic of mycoplasma species.

C,D,E & F. Electron micrographs of cells. Both WT and JCVI-syn1.0 show the same morphology

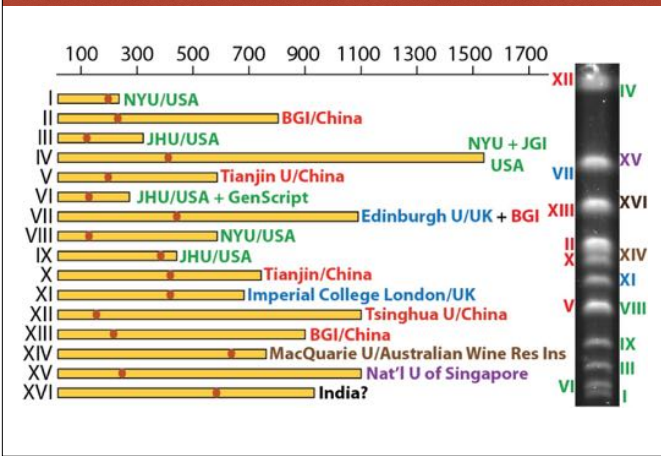


Synthetic bacterial chromosomes could then be transplanted into cells of related bacterial species after the destruction of the endogenous genome. Thus *Mycoplasma capricolum* could be converted to a synthetic version of *Mycoplasma mycoides*.

SC2.0 Synthetic Yeast Genome Project

31

Work is now underway on a much larger project: to create synthetic versions of the yeast (*Saccharomyces cerevisiae*) genome. This is an international project where individual chromosomes have been parcelled out to different institutions.



32

In a similar approach to the artificial bacterial chromosomes, small synthetic oligonucleotides are successively pieced together to create larger DNA fragments which in turn are progressively assembled into larger fragments or directly recombined into the target chromosome. Using alternative selection markers linear stretches of use chromosomes are progressively converted to the synthetic version. The refactoring of the 12 megabase yeast genome is largely complete. (<http://www.syntheticyeastresource.com>)

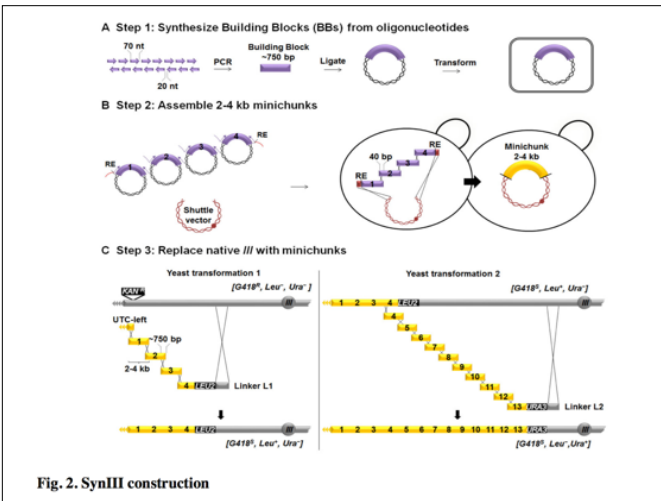
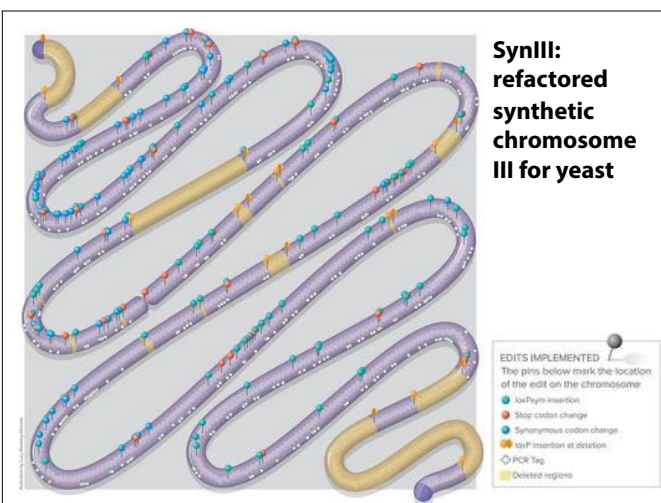


Fig. 2. SynIII construction

33

Schematic diagram of the refactored yeast chromosome III - showing (i) the introduction of *Lox* sites for scrambling the genome, (ii) altered codon usage, (iii) introduction of specific PCR tags and (iv) deletion of non-essential regions. The new DNA assembly techniques allow the possibility of building large-scale synthetic genetic circuits. The field faces the next challenge of integrating synthetic circuits with existing regulatory systems. We'll look at an example where a plant metabolic network has been integrated into microbe.





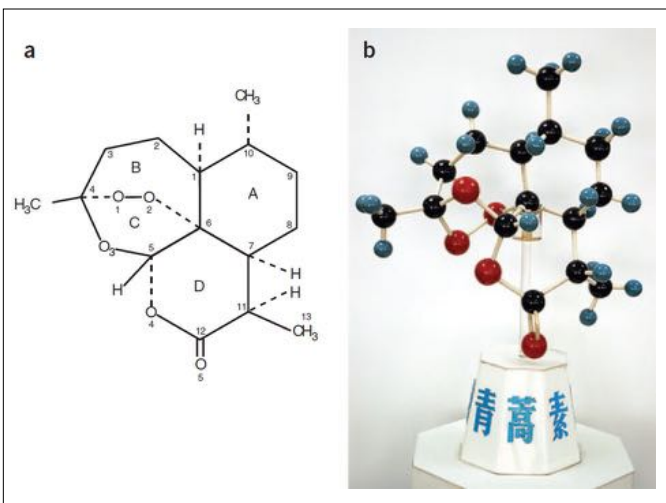
34

Mosquito borne malarial parasites have global distribution and affect millions. Resistance to malarial therapies occurs rapidly. Over time there has been a successive introduction of new therapies, as older therapies become less effective. The plant-derived drug artemisinin is a key component of current antimalarial therapies. Image of a mature *Artemisia annua* plant, and ornamental shrub from the daisy family (Asteraceae).



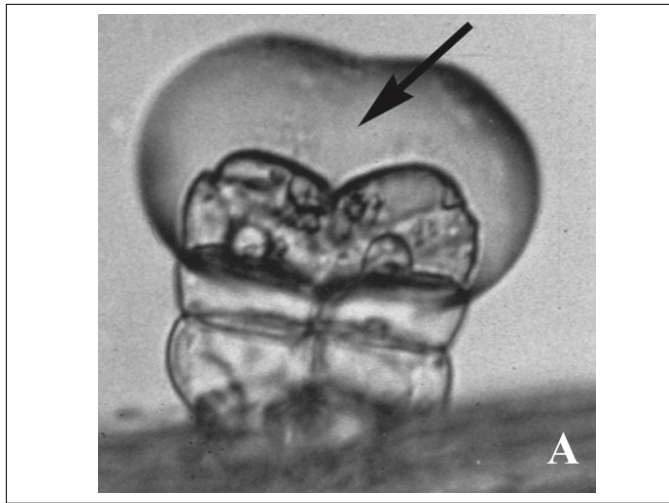
35

Chinese scientist Youyou Tu, who rediscovered the use of *Artemisia* extracts as an anti fever agent, and extracted the sesquiterpene artemisinin as an active anti-malarial drug. Key reference to production of active extracts from *Artemisia* (qinghao) plants from A Handbook of Prescriptions for Emergencies by Ge Hong (284–346 CE).



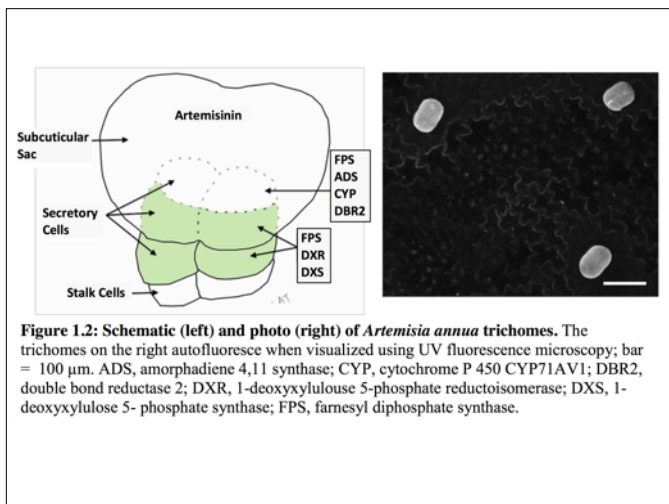
36

Artemisinin. (a) Molecular structure of artemisinin. (b) A three-dimensional model of artemisinin. Carbon atoms are represented by black balls, hydrogen atoms are blue and oxygen atoms are red. The Chinese characters underneath the model read Qinghaosu. The Chinese name for *Artemisia* is qinghao, and su means “basic element”.



37

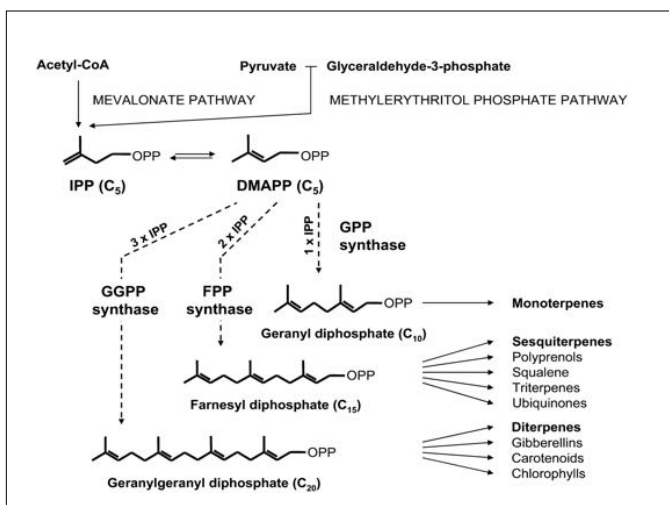
Micrograph of the surface of an *Artemisia* leaf. The trichomes, or leaf hairs, are the major source of artemisinin in the plant. The trichomes consist of a multicellular column which is highly active in the biosynthesis of artemisinin. This is capped by a subcuticular reservoir (arrowed) containing stored secondary compounds.



38

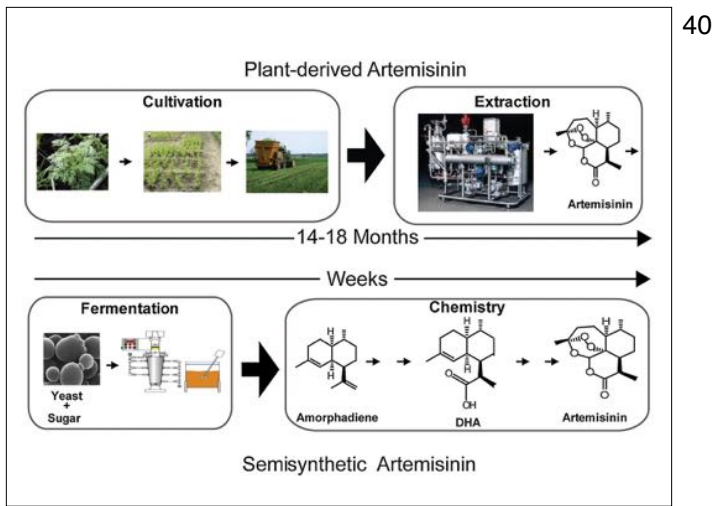
Localisation of biosynthetic enzymes in *Artemisia annua* trichomes. The column of cells express high levels of enzymes involved in the biosynthetic pathway. Products are secreted into the subcuticular sac. The cell complex behaves like a miniature biofactory.

Figure 1.2: Schematic (left) and photo (right) of *Artemisia annua* trichomes. The trichomes on the right autofluoresce when visualized using UV fluorescence microscopy; bar = 100 μ m. ADS, amorphaadiene 4,11 synthase; CYP, cytochrome P 450 CYP71AV1; DBR2, double bond reductase 2; DXR, 1-deoxyxylulose 5-phosphate reductoisomerase; DXS, 1-deoxyxylulose 5-phosphate synthase; FPS, farnesyl diphosphate synthase.

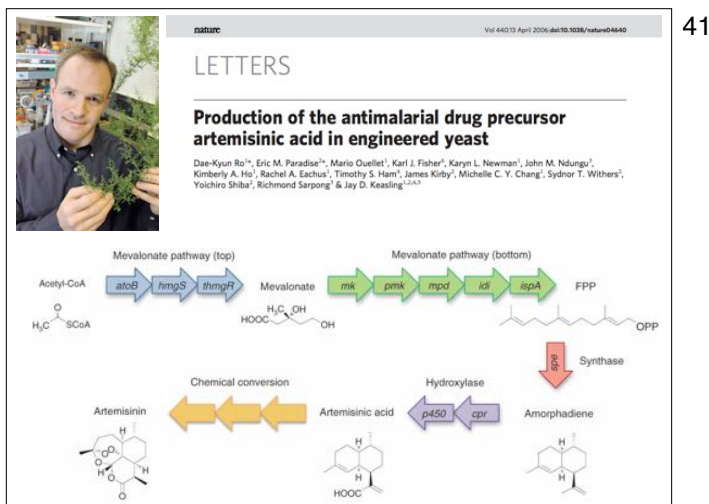


39

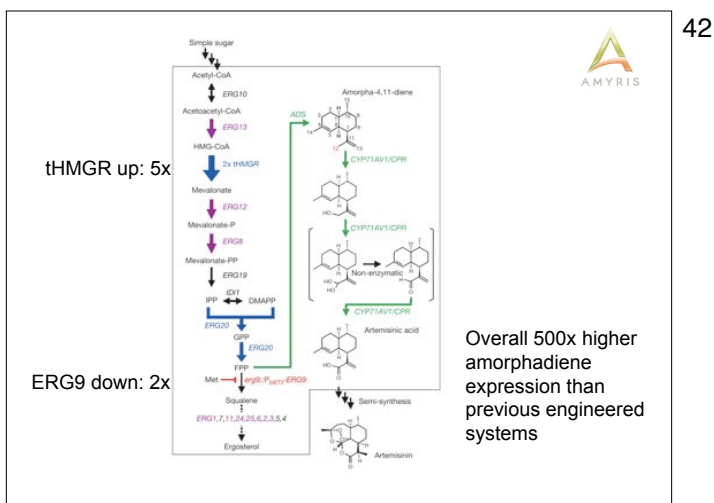
Artemisin is a sesquiterpene, member of the terpenoid family. The different classes of terpenes are synthesised by addition of different numbers of isoprene "units" and decorated by modifying enzymes. Farnesyl diphosphate (FPP) is the immediate precursor to artemisinin production in the cell.



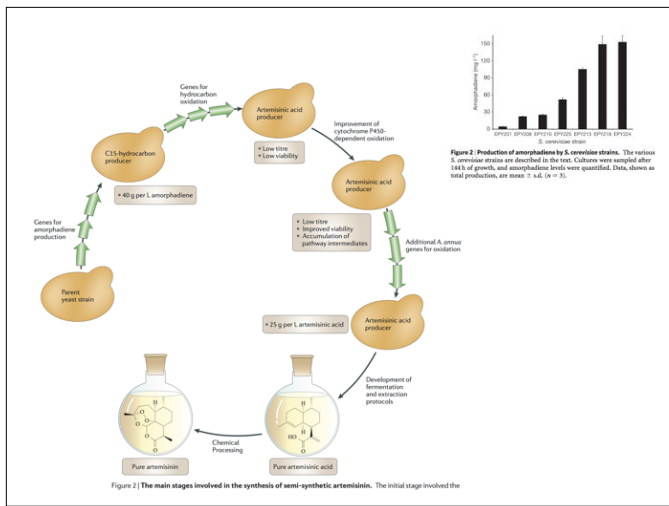
Artemisinin is naturally produced at low yields in slow-growing plants. Synthetic Biologists have taken up the challenge of transferring the artemisin pathway into yeast. This has potential benefits for lower cost and faster production of the drug.



Assembly of a hybrid pathway for artemisin production in yeast.



Schematic representation of the engineered artemisinic acid biosynthetic pathway in *S. cerevisiae*. Genes from the mevalonate pathway in *S. cerevisiae* that are directly upregulated are shown in blue; those that are indirectly upregulated by *upc2-1* expression are in purple; and the red line denotes repression of *ERG9* in strain EPY224. The pathway intermediates IPP, DMAPP and GPP are defined as isopentenyl pyrophosphate, dimethyl allyl pyrophosphate and geranyl pyrophosphate, respectively. Green arrows indicate the biochemical pathway leading from farnesyl pyrophosphate (FPP) to artemisinic acid, which was introduced into *S. cerevisiae* from *A. annua*. The three oxidation steps converting amorphadiene to artemisinic acid by *CYP71AV1* and *CPR* are shown.



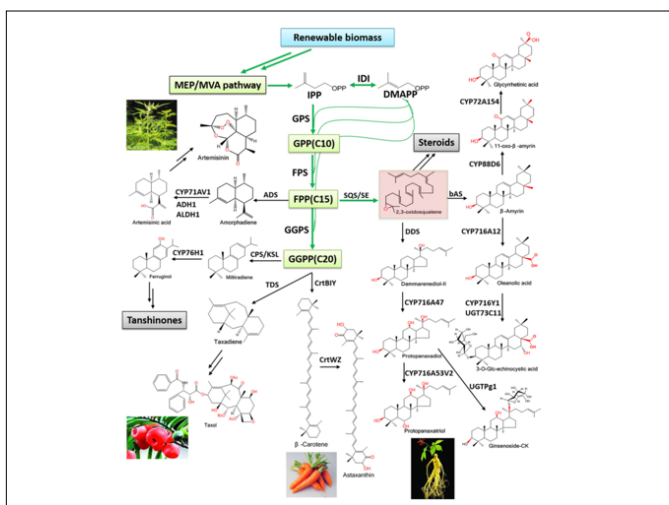
43

Representation of the main stages for improvement of production of semi-synthetic artemisin in yeast. Stepwise improvement of yields for artemisinin in engineered yeast strains.



44

Cooperative efforts required for discovery, research and development and production. Commercial production facility for semi-synthetic artemisinin, built by Sanofi in Italy.



45

A wide variety of secondary compounds derived from plants are potential candidates for microbial production.

Lecture 2

- 1. Microbial test systems**
- 2. Feedback-mediated regulation**
- 3. Repressilator**
- 4. Circuit optimisation**
- 5. *In vitro* systems for rapid testing**
- 6. Single cell analysis**
- 7. Computational modelling**
- 8. Megabase-scale genome refactoring**
- 9. Engineering of cellular metabolism**

Lecture 3: Reprogramming of multicellular systems.

Additional resources: <https://haseloff.plantsci.cam.ac.uk> (Education)

Genomics, Epigenetics & Synthetic Biology

Synthetic Biology and Plant Biotechnology: Lecture 3

Reprogramming of multicellular systems

Jim Haseloff

<https://haseloff.plantsci.cam.ac.uk>



1

Lecture 3

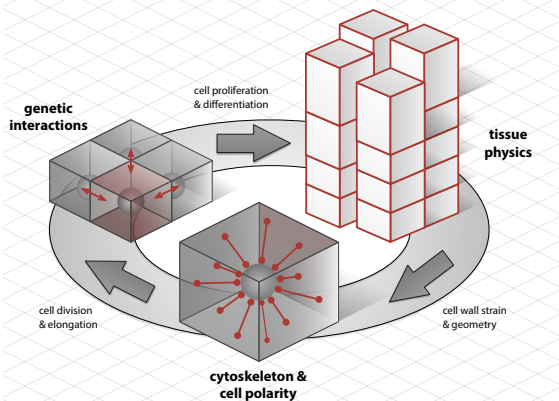
1. Multiscale organisation of tissues
2. Cellular interactions during growth
3. Wus-Clv interactions in meristem
4. Gene editing
5. Rewiring meristem interactions
6. Tomato fruit size
7. Inflorescence shape
8. Regulation of cell fates
9. Rewiring cell commitment

Additional resources: <https://haseloff.plantsci.cam.ac.uk> (Education)

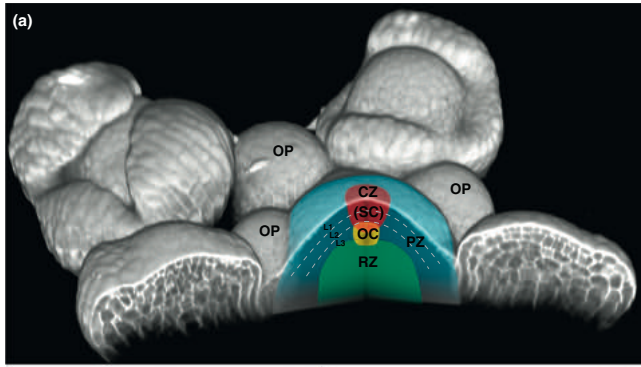
2

In this lecture, we will look at circuits which (i) act cell non cell autonomously, and (ii) work in a multicellular context and involve communication between cells.

3



Multi-scale view of plant growth. (i) Interaction between cytoskeletal elements and local cell wall determinants, strain or geometry regulates the polarity of cell division and elongation. (ii) Genetic interactions between neighbouring cells trigger gene expression, cell proliferation and differentiation. (iii) Cellular growth results in physical strains that are transmitted across tissues and constrain cell growth. (iv) Physical constraints on cell size and shape regulate timing and orientation of individual cell divisions and guide morphogenesis.

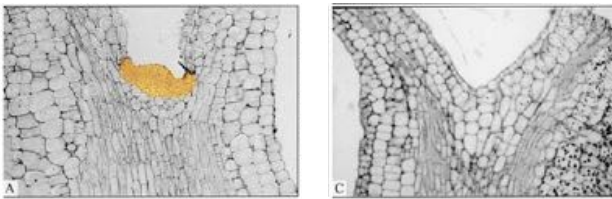


Engineering multicellular architectures in plants

4

In order to reprogram plant systems, it will be necessary to include cell cell communication. This will be required to initiate, maintain and coordinate patterns of gene expression in a multicellular context. We will look at an example of manipulation of shoot growth in a higher plant.

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

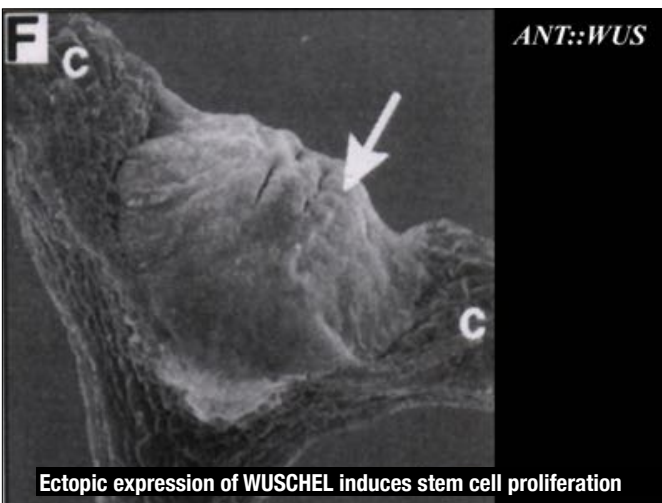


wt

wuschel
(*wus*)

5

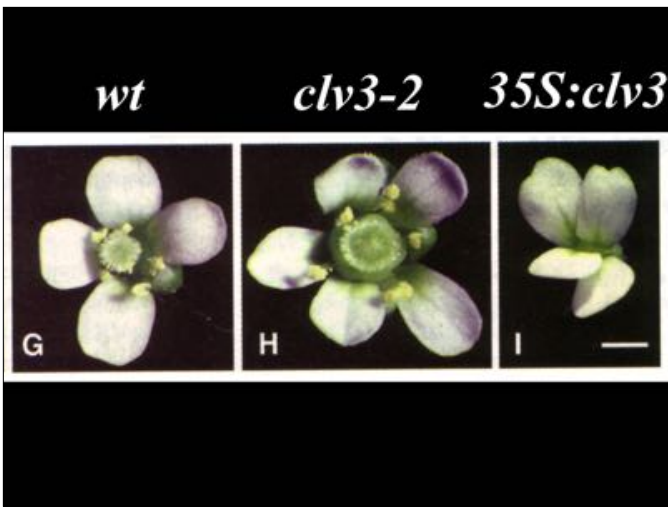
The growth of the apical meristem in high apical meristem in higher plants is controlled by opposing sets of regulatory proteins. We will look at two of these, as characterised in the model plant Arabidopsis. First, Wuschel is a homeodomain gene that is required for maintenance of the stem cells within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



6

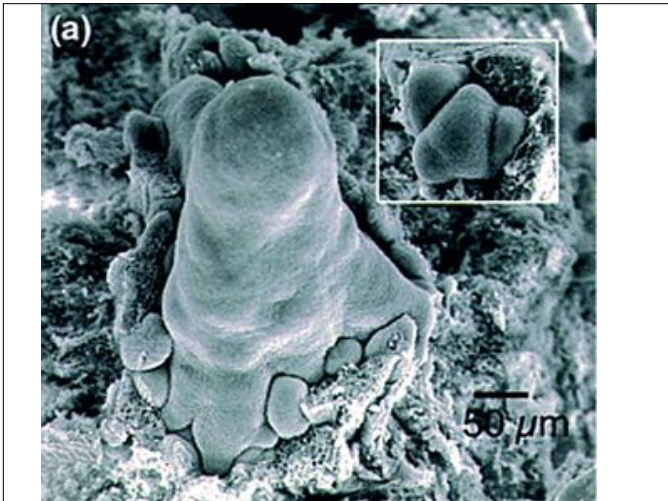
Overexpression of the Wuschel gene results in the hyper-proliferation of stem cells within the meristem. Wuschel is sufficient for the production of indeterminacy in targeted cells.

7



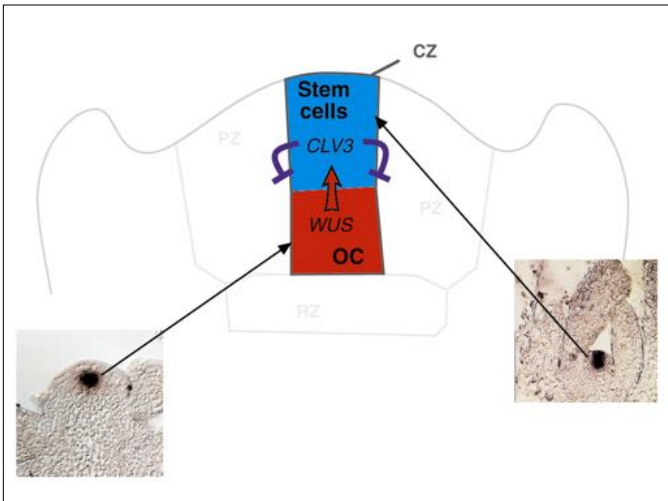
In contrast, the *Clavata3* gene plays a role in suppressing the size of the meristem. In its absence the meristem enlarges (which can result in the oversize *Arabidopsis* flowers shown here). Overexpression of *Clavata3* results in loss of meristem activity (and undersized flowers).

8

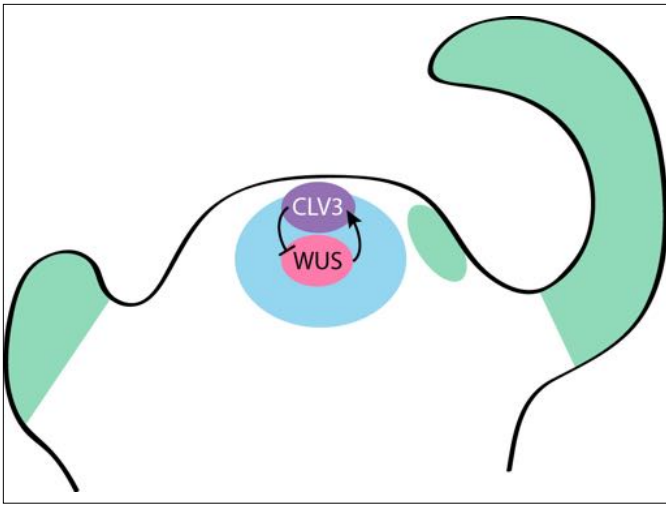


Scanning electron micrographs of wild-type (inset) and *clavata3* mutant meristems. Loss of a *clavata3* activity results in gross enlargement of this floral inflorescence meristem.

9

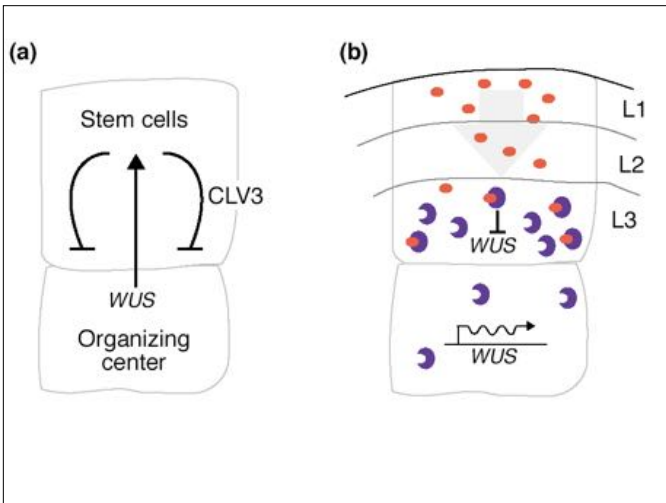


The expression of both *Wuschel* and *Clavata3* is tightly localised to the shoot apex of the *Arabidopsis* plant. They possess overlapping expression patterns, with *Clavata3* expression positioned to the outside of the meristem. *Wuschel* expression is localised inside the meristem.



10

Wuschel and Clavata3 form part of a feedback loop that regulates meristem size in Arabidopsis. Wuschel gene activity promotes cell proliferation, and Clavata3 acts in a negative feedback loop to suppress meristem growth.



11

Clavata3 encodes a peptide that is produced on the outside of the meristem, diffuses to the cells on the inside, binds to receptors on these cells, and inhibits Wuschel expression via the MAP kinase pathway. These mechanisms of action have been discovered through work with the model plant Arabidopsis. However, the same mechanisms are at work in most other higher plants.

CRISPR, THE DISRUPTOR

A powerful gene-editing technology is the biggest game changer to hit biology since PCR. But with its huge potential come pressing concerns.

Introduction/revision for CRISPR-Cas9 gene editing

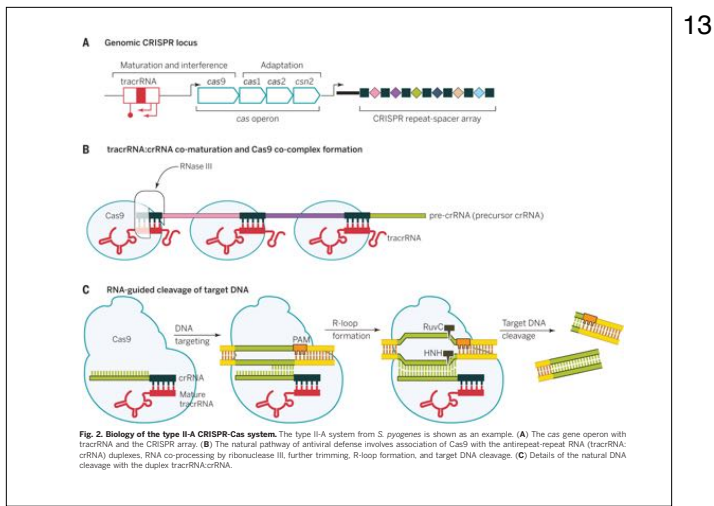
RIDING THE CRISPR WAVE

Biologists are embracing the power of gene-editing tools to explore genomes.

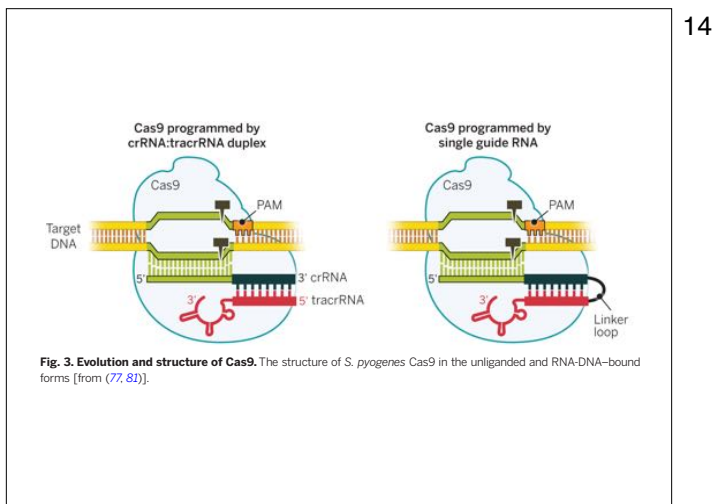
CRISPR DESTINIES

12

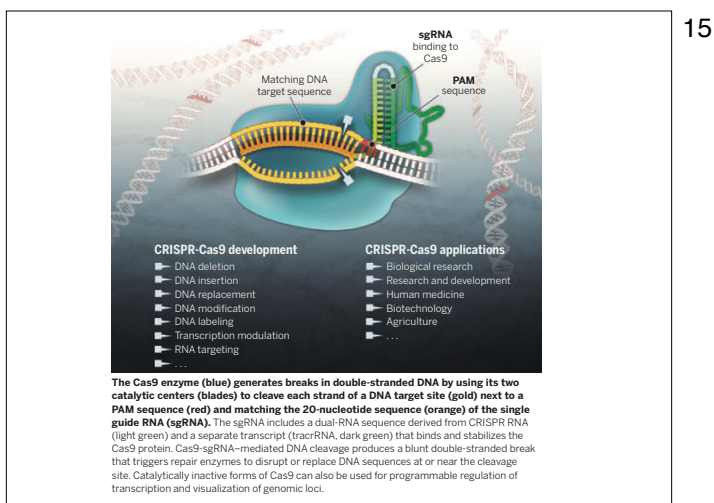
There has been an explosion of new gene editing techniques and their application for biomedical and agricultural uses. Accessible articles and reviews that describe the new wave of editing techniques can be found with other lecture materials at <http://www.haseloff-lab.org/education>



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.

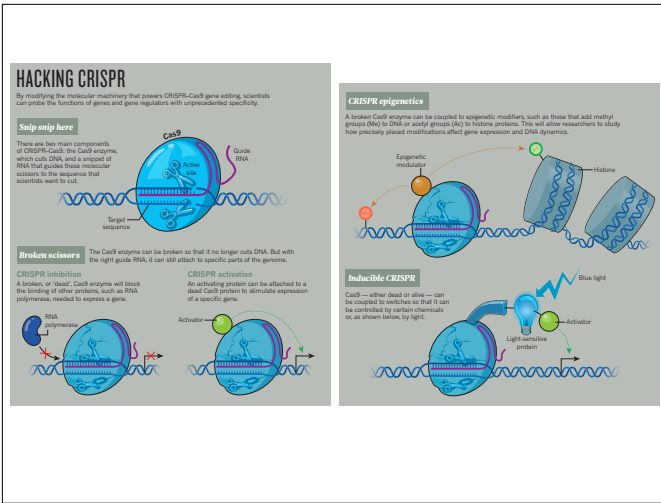


The *tracrRNA* and *crRNA* components of the nuclease can be fused to create a single guide sequence that, in combination with Cas9, will produce a nuclease that can be targeted to any DNA sequence adjacent to a 3 nucleotide PAM sequence.

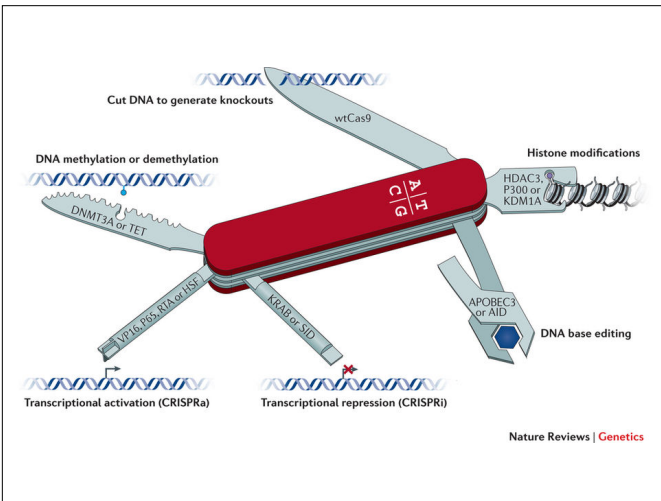


The CRISPR-Cas9 system can be used to create a programmable DNA binding complex. This will normally create a double-strand break at the target site. This has been used widely for targeted mutagenesis, via error-prone repair of dsDNA breaks *in vivo*, and to promote DNA replacement through homologous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

Use of the CRISPR-Cas9 complex to catalyse dsDNA breaks, site-specific delivery of inhibitor or activators of transcription, recruitment of chromatin modifiers, and as an inducible (and targeted) regulator of gene expression.



The RNA-programmable DNA binding element is finding many applications as a tool for genome manipulation *in vivo*.



History of CRISPR-Cas9 manipulation and commercial exploitation.

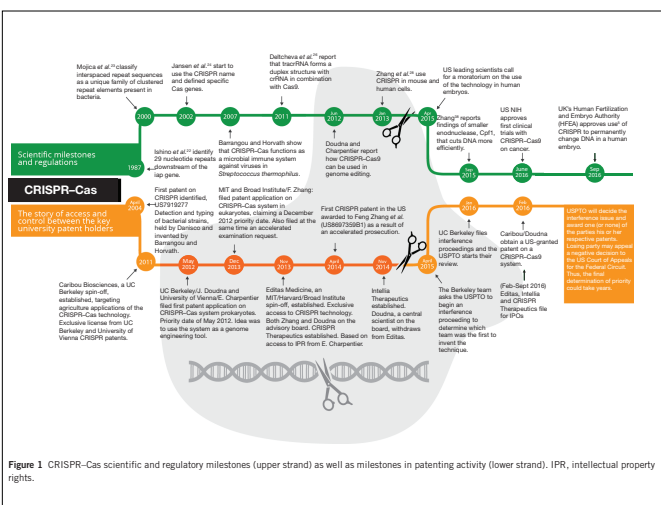
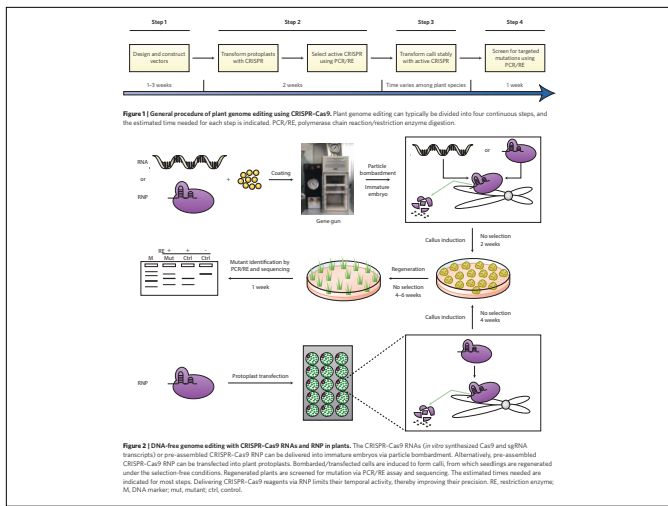
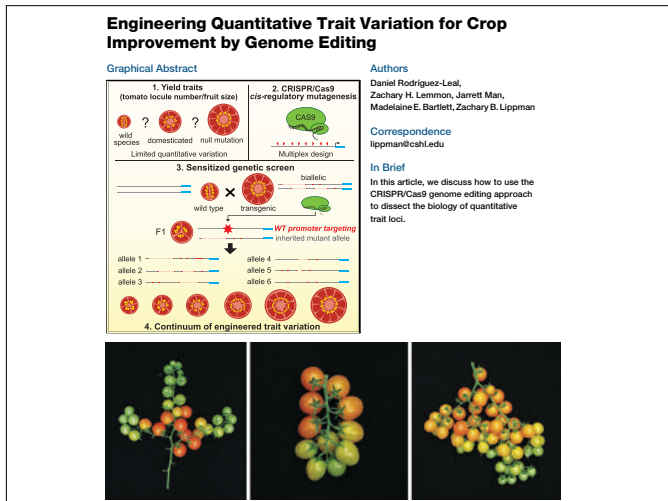


Figure 1 CRISPR-Cas scientific and regulatory milestones (upper strand) as well as milestones in patenting activity (lower strand). IPR, intellectual property rights.



19

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.



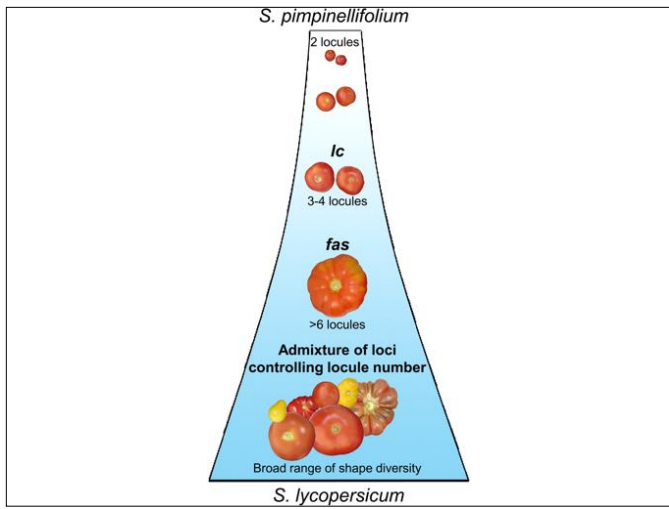
20

In a recently published experiment, Lippman and colleagues targeted regulatory elements in the tomato genome, using CRISPR-Cas9 delivery. They could generate variant traits in a targeted way, and produce plant lines with traits that could be introduced directly into a breeding programme. This is demonstration of an alternative to conventional plant transformation, and introduction of foreign activities - that has been potential to be regulated differently from existing GM crop systems.



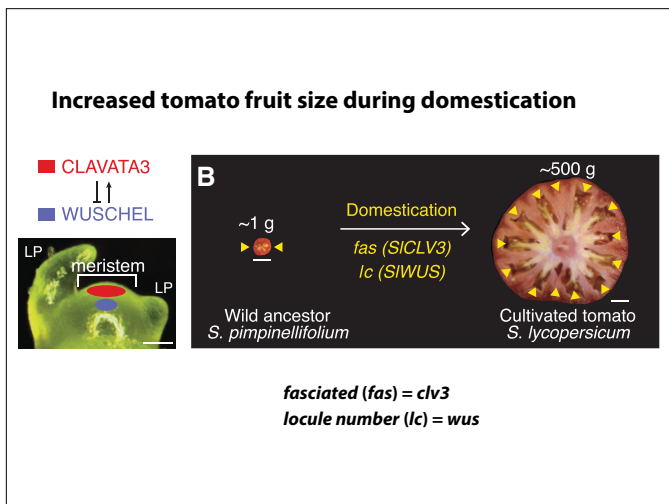
21

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.



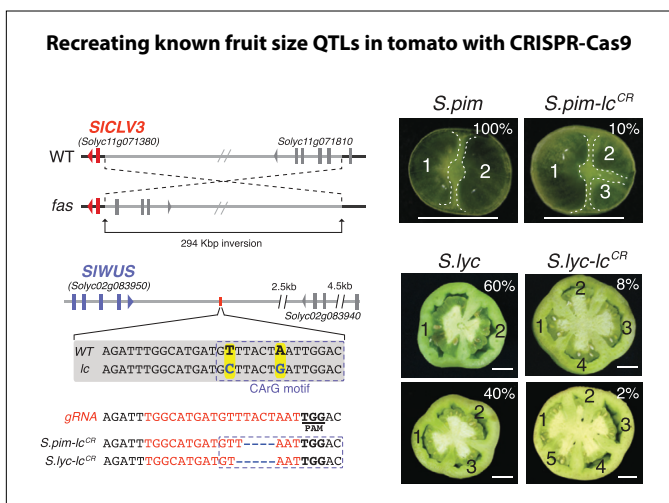
22

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* (*fc*).



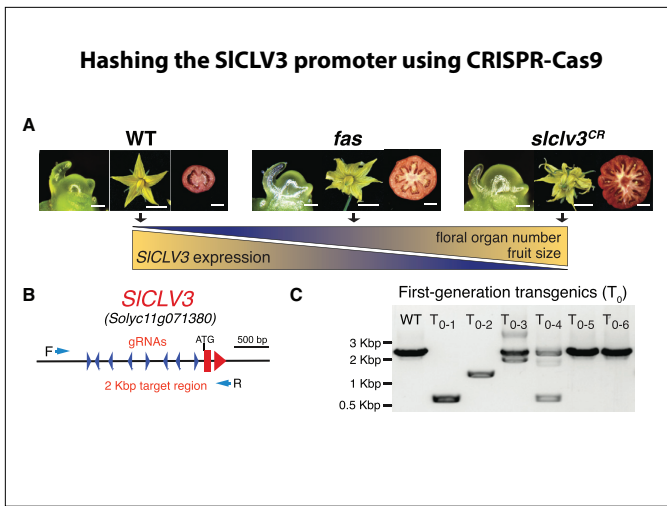
23

Fas encodes *Solanum lycopersicum* CLV3, and *Lc* encodes WUS. The conserved CLV3-WUS negative feedback circuit controls meristem and fruit size in tomato. (LP, leaf primordia). (B) The fasciated (*fas*) and locule number (*lc*) fruit size QTL both contributed to increased tomato fruit size and locule number during domestication. Yellow arrowheads, locules.



24

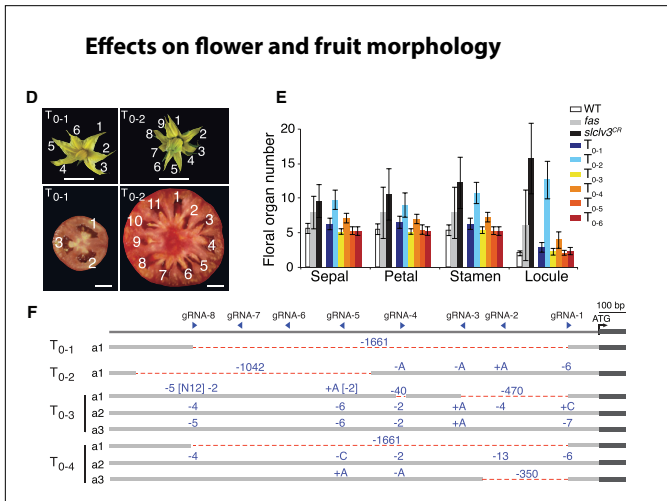
1. The *fas* mutation is caused by an inversion with a breakpoint 1 Kbp upstream of *SICLV3*. (D) The *lc* QTL (red rectangle) is associated with two SNPs (in bold) in a putative repressor motif (CArG, blue-dashed square) 1.7 Kbp downstream of *SIWUS*.
2. CRISPR/Cas9-induced deletions in the CArG repressor motif (blue-dashed square) of *Solanum pimpinellifolium* (*S. pim*) and *Solanum lycopersicum* (*S. lyc*). The gRNA target sequence is highlighted in red and the PAM site underlined.
3. *S.pim-lc^{CR}* plants produce fruits with more than two locules. *S.pim-fasNIL S.pim-lc^{CR}* double mutants synergistically increase locule number.



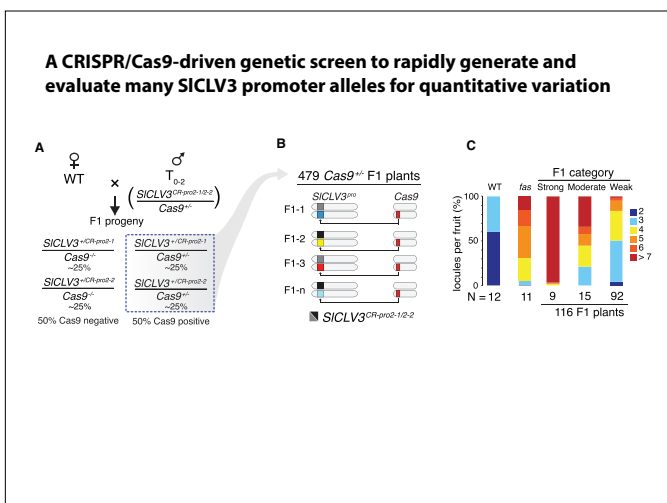
A. Model showing how an allelic series of SICLV3 transcriptional alleles could provide a range of quantitative effects on floral organ number according to a simple linear relationship of reduced expression resulting in increased phenotypic severity. (WT, fas, and clv3CR are shown as reference points in this hypothetical continuous relationship).

B. Schematic of SICLV3 promoter targeted by eight gRNAs (numbered blue arrowheads). Blue arrows, PCR primers.

C. PCR showing multiple deletion alleles in four T0 plants. Amplicons were obtained using primers spanning the entire target region.



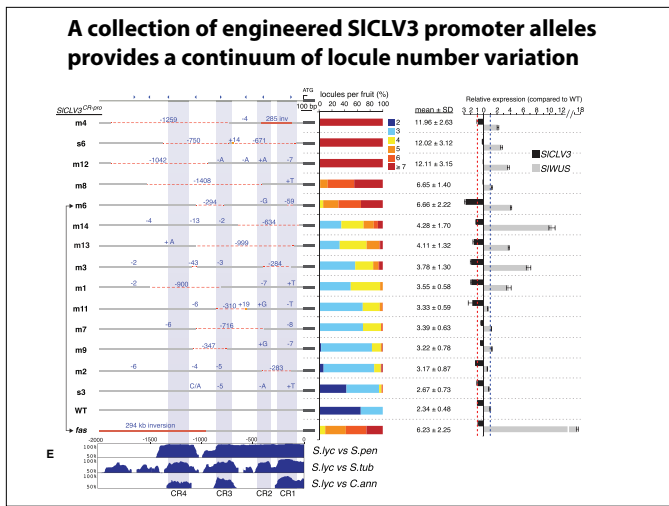
Weak and strong effects on flower morphology and fruit size were observed among T0 lines. Number of floral organs and locules are indicated. (E) Quantification of floral organ number (mean \pm SD; $n > 10$) in T0, WT, fas, and slclv3CR plants. (F) Sequencing of SICLV3 promoter alleles for all T0 plants. Deletions (-) and insertions (+) indicated by numbers or letters. T0-5 and T0-6 contained only WT alleles (data not shown). Blue arrowheads, gRNAs; a, allele.



A. Crossing scheme for generating a sensitized F1 population heterozygous for an inherited allele and segregating for a CRISPR/Cas9 transgene (blue-dashed square). Expected segregation frequencies are indicated (%).

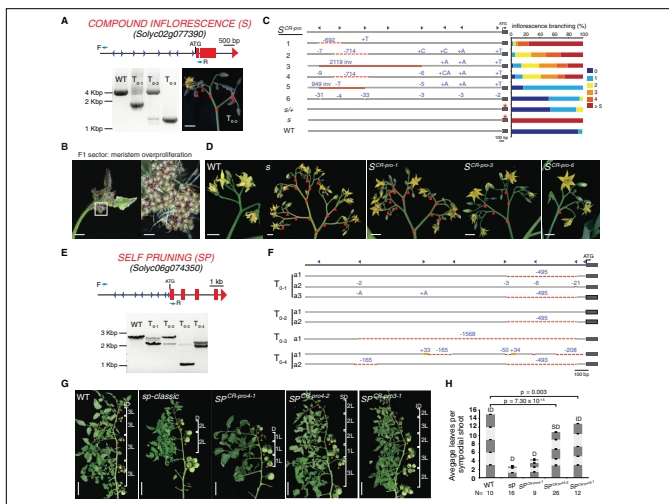
B. Model showing how Cas9 activity in Cas9+ hemizygous plants creates new mutant alleles (colored boxes) by targeting the WT SICLV3 promoter (SICLV3pro) introduced from the cross. Alleles derived from T0-2 are shown as black or dark gray boxes. The transgene containing the CRISPR/Cas9 cassette (Cas9) is shown (red box).

C. Locule number for WT, fas, and F1 plants grouped into three phenotypic categories: strong, moderate, weak.



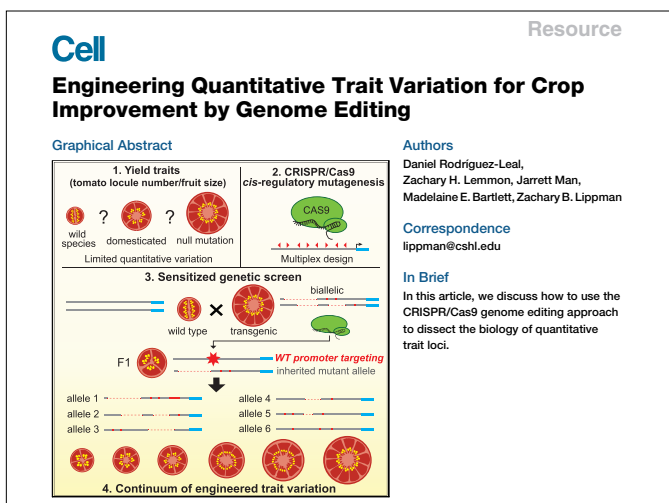
28

- Sequences of 14 new SICLV3pro alleles. Deletions (-) and insertions (+) indicated as numbers or letters. gRNAs, blue arrowheads. Parental F1s marked at right.
- qRT-PCR of SICLV3 and SIWUS from reproductive meristems for WT, fas, and each SICLV3CR-pro allele. Dashed lines mark WT levels for SICLV3 (red) and SIWUS (blue)



29

CRISPR/Cas9 Targeting of the Compound Inflorescence (S) and Self Pruning (SP) promoters results in a further range of modified inflorescence and plant architectures.



30

- CRISPR/Cas9 targeting of a cis-regulatory motif recreated a domestication QTL.
- CRISPR/Cas9 drove mutagenesis of promoters to create a continuum of variation.
- Phenotypic effects were not predictable from allele type or transcriptional change.
- Selected promoter alleles in developmental genes could improve yield traits.

31

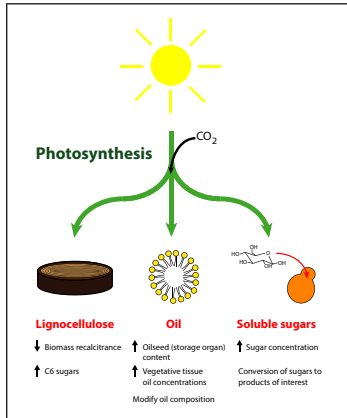
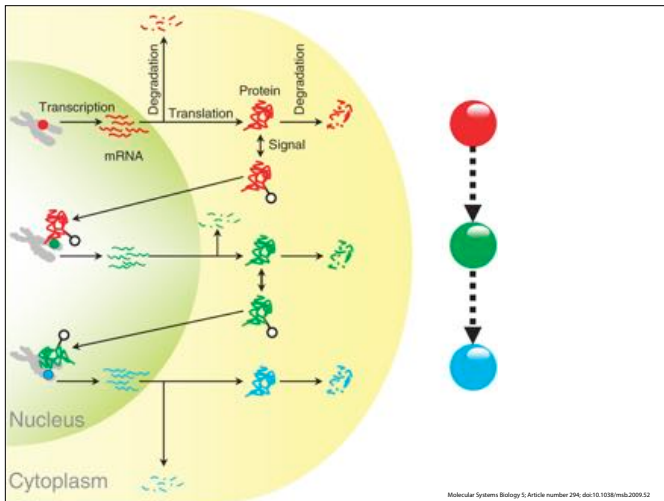


Figure 2. Application of synthetic biology to improve and engineer bioenergy traits.

The Plant Journal, (2016), 87, 103–117

Can one also manipulate regulatory proteins in order to manipulate plant products that are produced by metabolic networks?

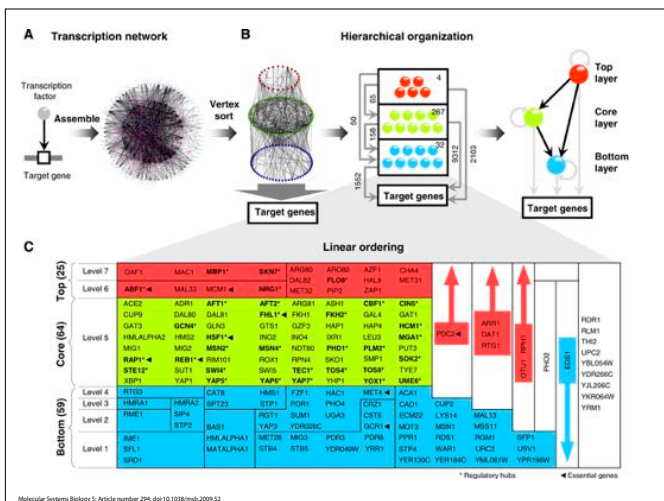
32



Molecular Systems Biology 5, Article number 294, doi:10.1038/msb.2009.52

Transcriptional cascades underpin gene regulatory networks. Transcription factors (TFs), denoted as nodes in a network (red and green circles), represent several entities (gene, mRNA, and protein) and events (transcription, translation, degradation, etc) that are compressed in both space and time. The series of regulatory events can be conveniently represented as a node in the network, although this does not capture the dynamics of these entities and the biological processes.

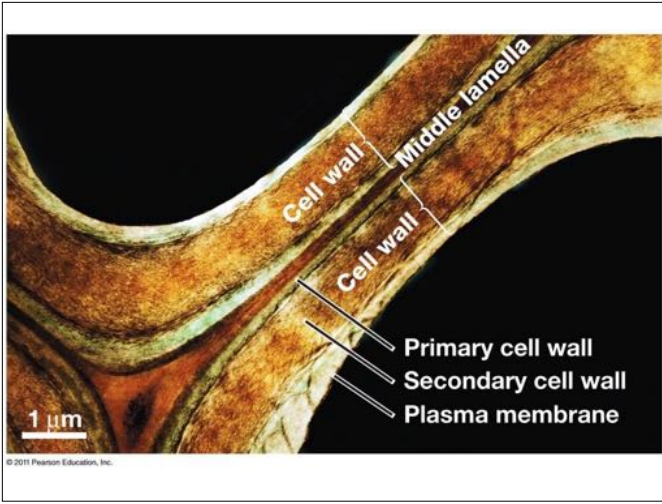
33



Molecular Systems Biology 5, Article number 294, doi:10.1038/msb.2009.52

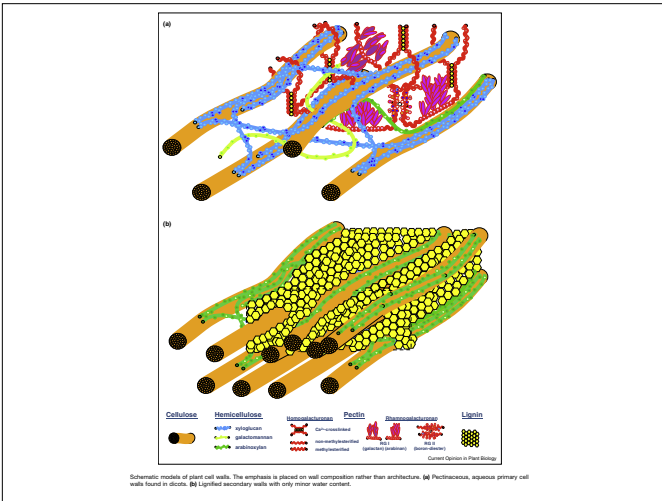
An example: hierarchical structure in the yeast transcription regulatory network. The organization of transcription factors in the network naturally clusters into three basic non-overlapping layers: the top (red), core (green), and bottom (blue). Thirty-two regulatory hubs are highlighted in bold and marked with a star (*), and nine essential TFs are marked with an arrow.

34



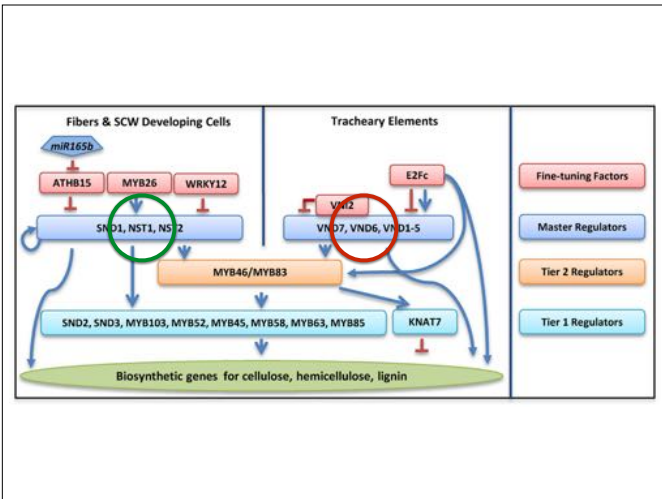
As an example, we will look at the engineering of transcription factors to modify cell wall composition in plants. The development of improved feedstocks for bioenergy production has focused on maximising primary cell wall and cellulosic material, and limiting amounts of lignin compounds during secondary wall growth.

35

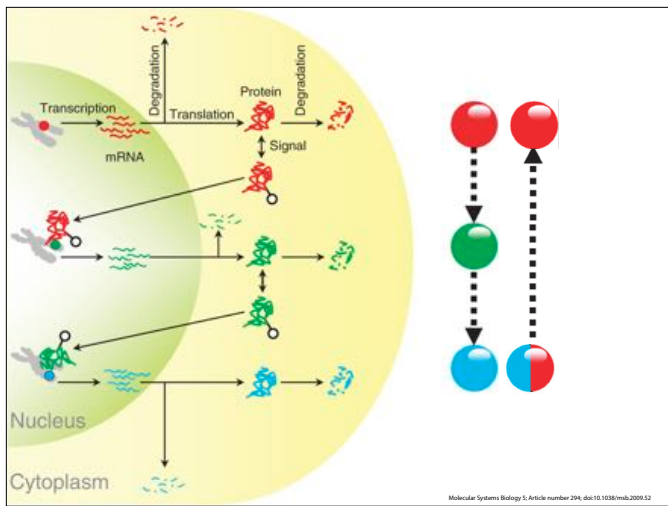


Primary cell walls are composed mainly of cellulose micro fibrils and cross-linked hemicellulose and pectins. During secondary wall deposition lignin is deposited which results in cell wall material that is chemically resistant, difficult to digest for biofuel production.

36

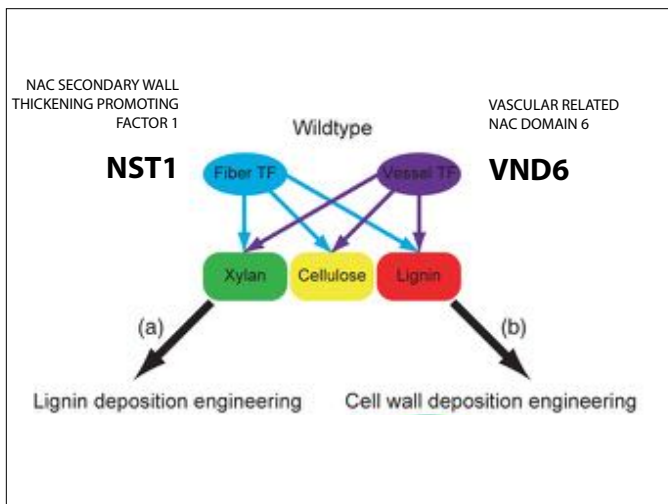


A hierarchical transcriptional network controls biosynthetic genes that are responsible for timing, synthesis and balance of different cell wall carbohydrates. This is best characterised in *Arabidopsis*. Locque and colleagues have developed a technique for reprogramming this transcriptional hierarchy - and increasing the balance of resources devoted to (i) fibres and secondary cell wall material vs. (ii) lignified tracheary elements.



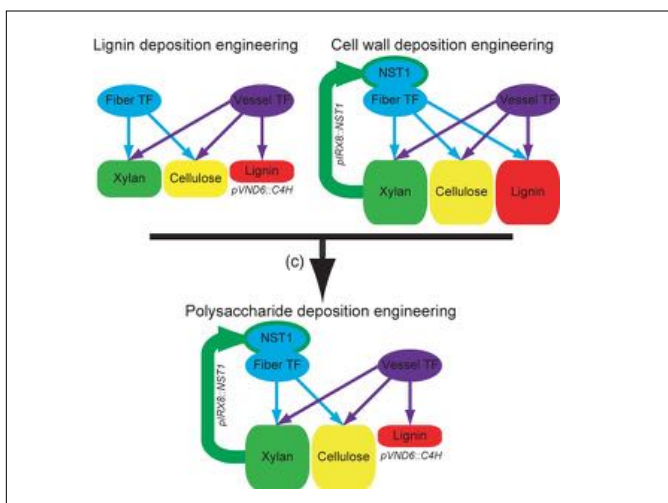
37

In this scheme, a gene from downstream in the regulatory hierarchy is identified. Its promoter (blue) is fused to a copy of the master regulator (red). Therefore when the master regulator is switched on during the normal course of development, there is a cascade of transcriptional control events that results in triggering of housekeeping genes, and in addition, triggering of a new copy of the master regulator. Results in expansion of the normal pathway.



38

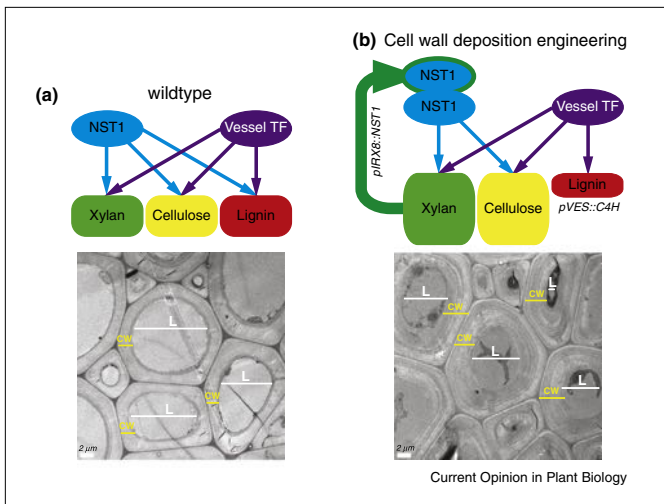
The two "arms" of vascular cell development are regulated by two transcription regulators, NAC Secondary Wall Thickening Promoting Factor 1 (NST1) and Vascular Related NAC Domain 6 (VND6). These genes show expression patterns that are limited to fibre and vessel cells, respectively.



39

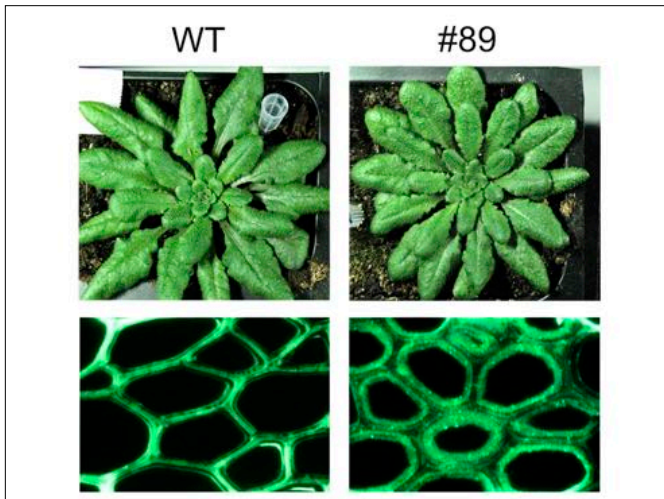
In a two-way approach, Loque and colleagues interfered with the synthesis and deposition of lignin. (i) The promoter of a key lignin gene, *C4H*, was replaced by the vessel-specific promoter of transcription factor *VND6* in a *c4h* mutant. This rewired lignin biosynthesis specifically for vessel formation while disconnecting *C4H* expression from the fibre regulatory network. (ii) The promoter of the *IRX8* gene, a secondary cell wall glycosyltransferase, was used to express a new copy of the fibre transcription factor *NST1*, and as the *IRX8* promoter is induced by *NST1*, this creates an artificial positive feedback loop

40



Engineering of increased cell wall density and decreased lignin in Arabidopsis. Schematic of simplified regulatory network controlling secondary cell wall biosynthesis in vessel and fibre cells in plants and images from wildtype (a) and engineered (b) Arabidopsis plants depicting interfascicular tissues composed of fiber cells. Engineered plants were generated from a *c4h* defective mutant (mutant affected in the second lignin biosynthesis step) that was transformed with the wild-type version of the mutated *C4H* gene driven by a vessel-specific promoter which rescued the negative effect of low lignin content. Generated plants were further transformed with a construct (*pIRX8::NST1*) that led to higher expression of master transcription factor controlling secondary cell wall biosynthesis in fiber cells (e.g. *NST1*).

41



Production of transgenic plants with increased xylan content and decreased lignin content.

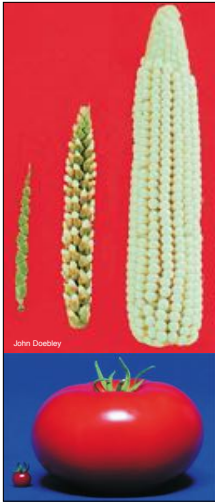
42

Lecture 3

1. Multiscale organisation of tissues
2. Cellular interactions during growth
3. Wus-Clv interactions in meristem
4. Gene editing
5. Rewiring meristem interactions
6. Tomato fruit size
7. Inflorescence shape
8. Regulation of cell fates
9. Rewiring cell commitment

Additional resources: <https://haseloff.plantsci.cam.ac.uk> (Education)

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.



Reprogramming plants

- Plants provide proven, global, low-cost technology for gigatonne scale bioproduction
- We need faster, simpler multicellular systems for engineering form and metabolism
- Synthetic biology offers breakout technologies