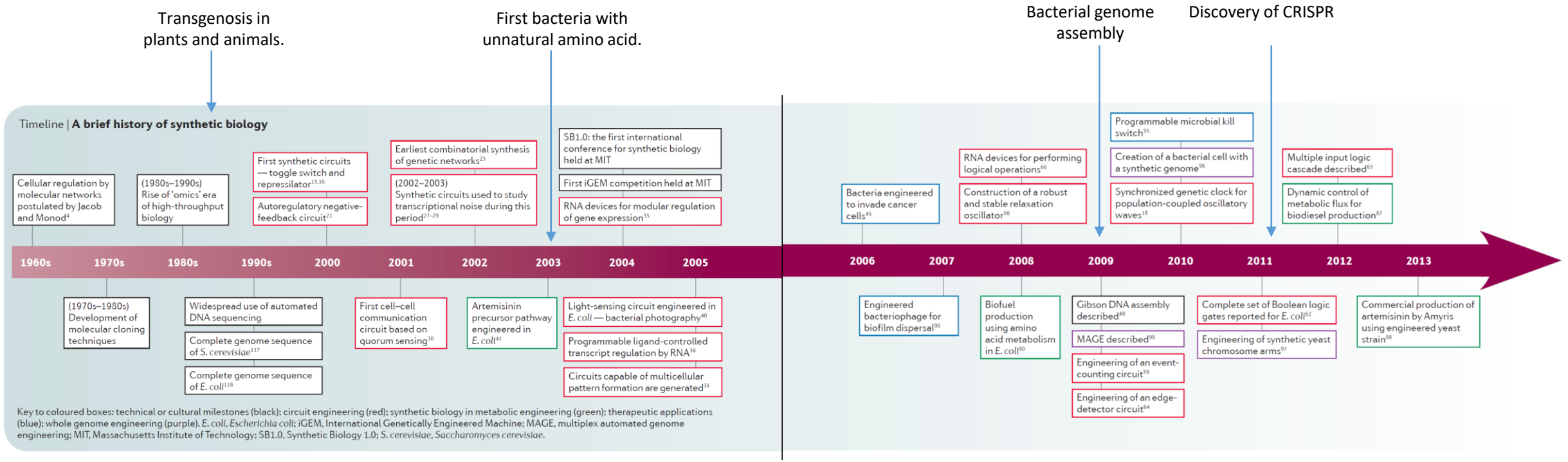
A brief history of synthetic biology



Cameron et al: Nat Rev Microbiol 2014

Synthetic Biology Market is Expected to Reach $38.7 Billion, Globally, by 2020 - Allied Market Research

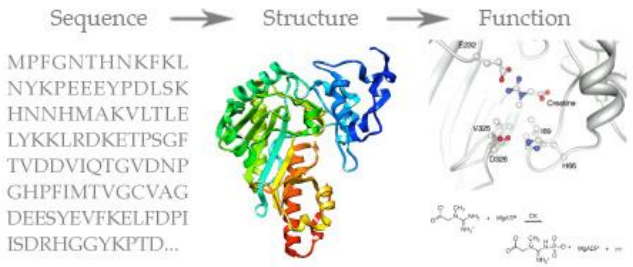
DNA synthesis is the largest segment within enabling products segment, whereas oligonucleotide synthesis is expected to be fastest growing market at 57.8% CAGR during 2014 and 2020. **Chassis organism** would be the fastest growing core product during the forecast period with synthetic DNA occupying largest market share. Other core products included in the study are **synthetic genes**, **synthetic scells**, and XNA. **Biofuels**, within enabled product segment, is expected to exhibit tremendous growth; registering a CAGR of 110.1% during forecast period. However, **synthetic biology-based pharmaceuticals and diagnostics products** will generate largest amount of revenue within enabled product segment followed by **agriculture** and **chemicals** sub-segments.

Hierarchical organization of living matter

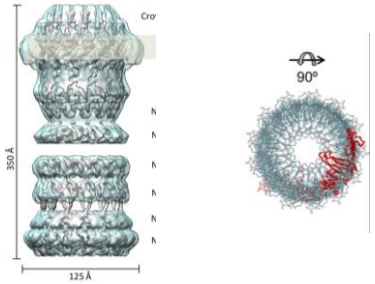
- the blueprint: DNA

  ...AACGGCTAATCTGG...

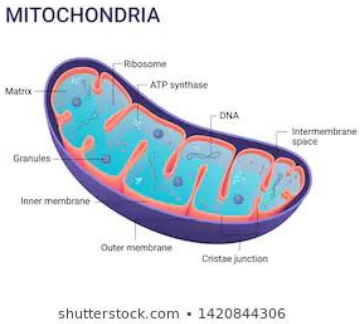
- the effectors: proteins



- nanomachines: protein complexes

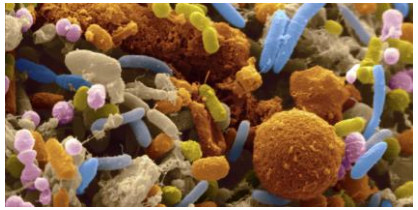


- cellular factories: cellular facotries



- autonomous units of life: cells

- organisms and ecosystems



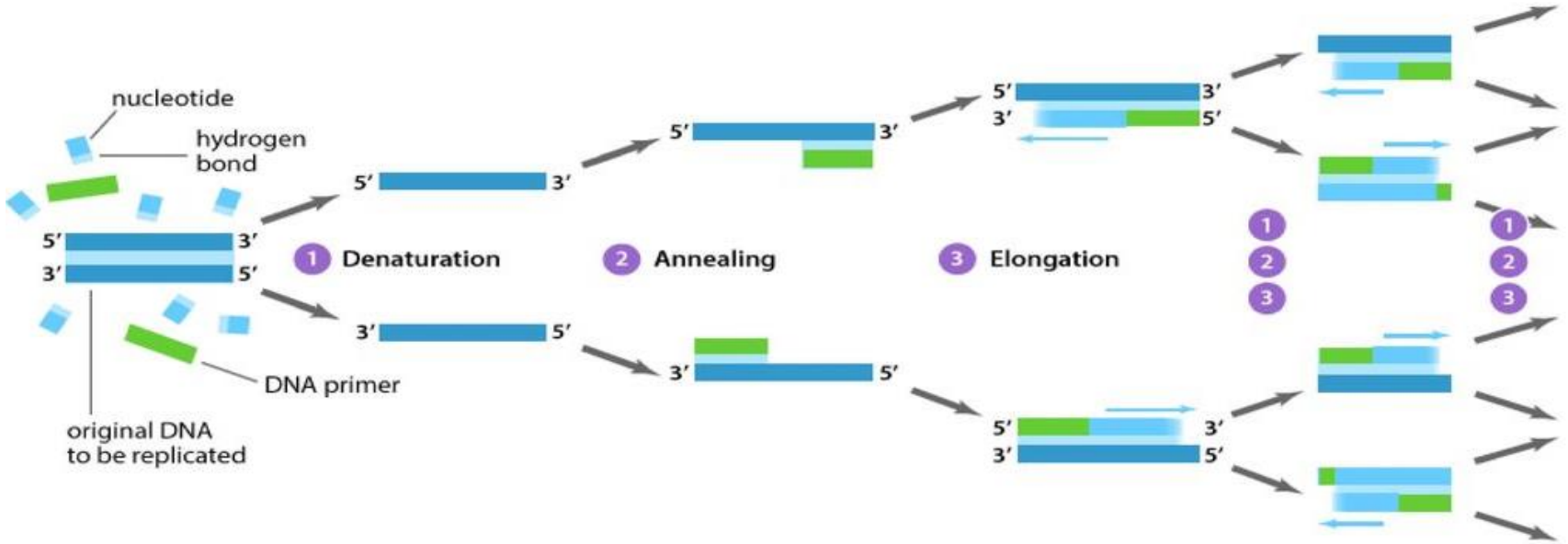
Top-down and botom-up approaches in synthetic biology

The top-down approach derives a synthetic cell from a biological cell by manipulating, for example, its genes and protein content. This type of synthetic cell is typically living and still closely related to its biological ancestor. Top down synthetic biology involves using metabolic and genetic engineering techniques to impart new functions to living cells.

The bottom-up approach, on the contrary, starts with nonliving matter. Its most basic synthetic cell is merely a cell-sized compartment. Cell-like functionality is derived by reconstituting functional modules, made from natural or artificial molecular building blocks. The complexity of the synthetic cell is increased step-by-step, by including more and more components. The ultimate goal is contraction of an artificial cell (reconstruction of life).

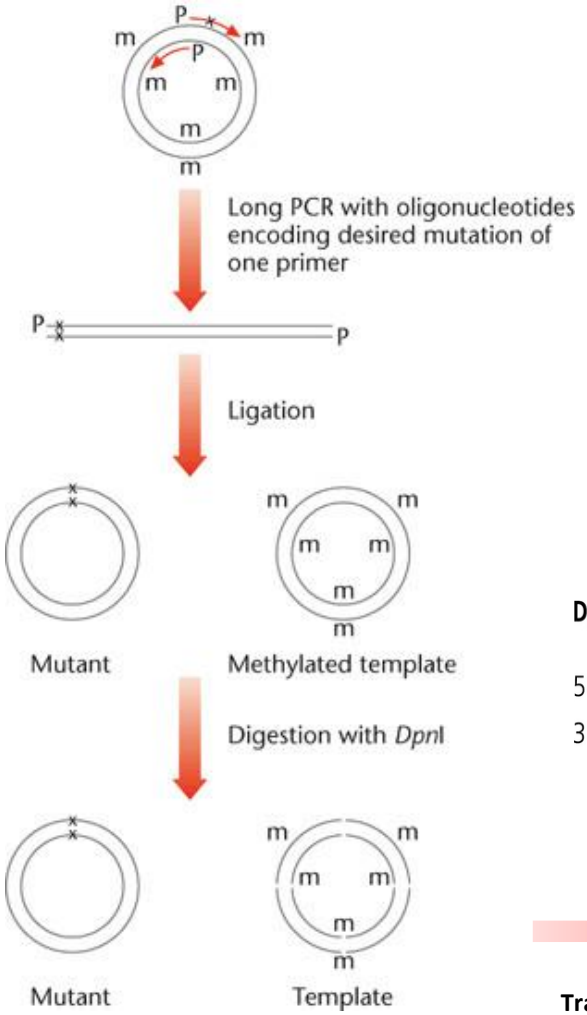
Enabling technologies: DNA assembly

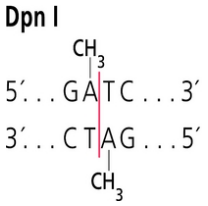
Polymerase chain reaction (PCR)



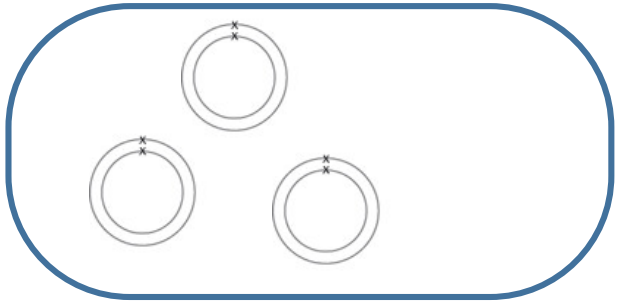
Amplification of up to 20 kbDNA fragment from pre-existing template (genomic DNA, cDNA library, cloned fragment)

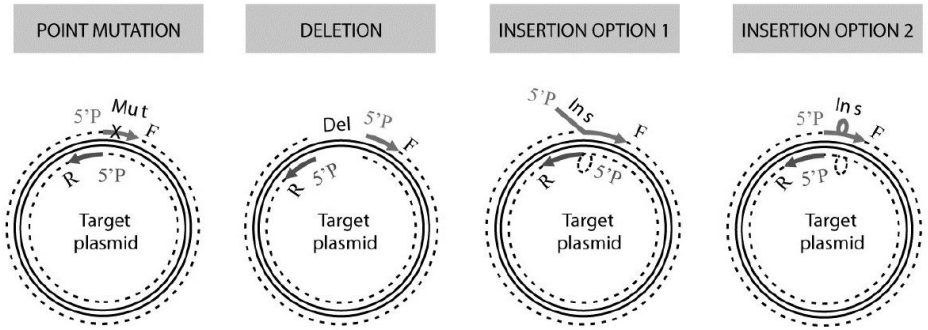
Site directed mutagenesis



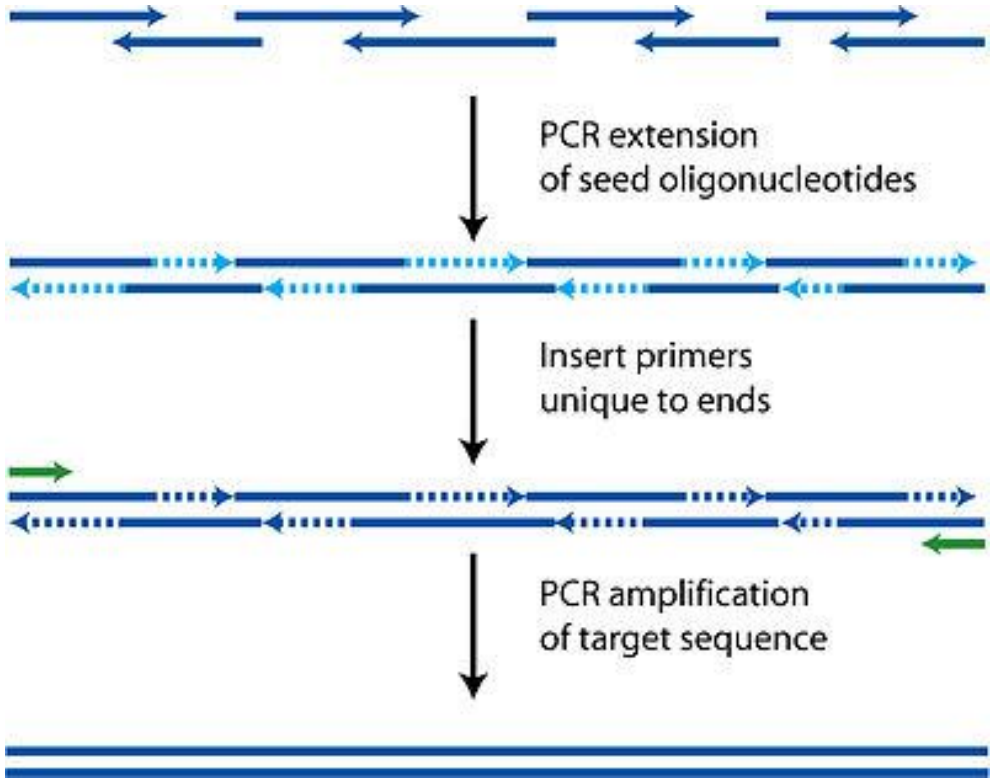


Transformation



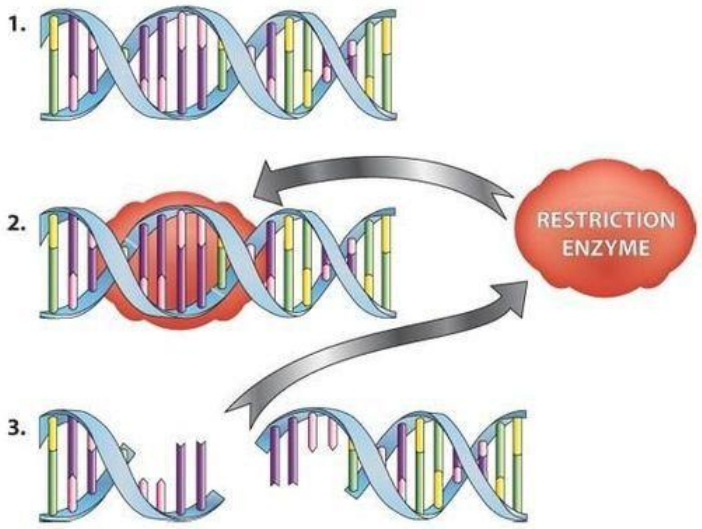


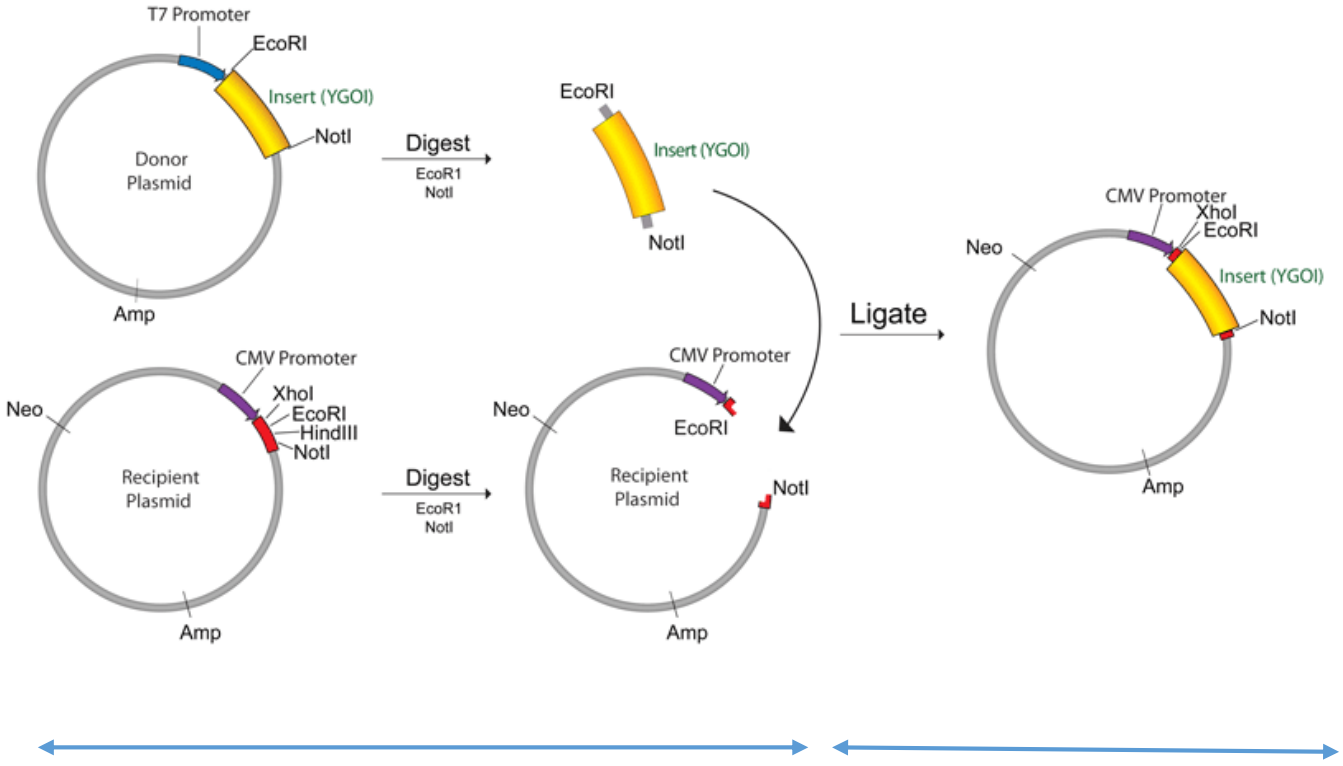
Synthetic genes: Polymerase chain assembly



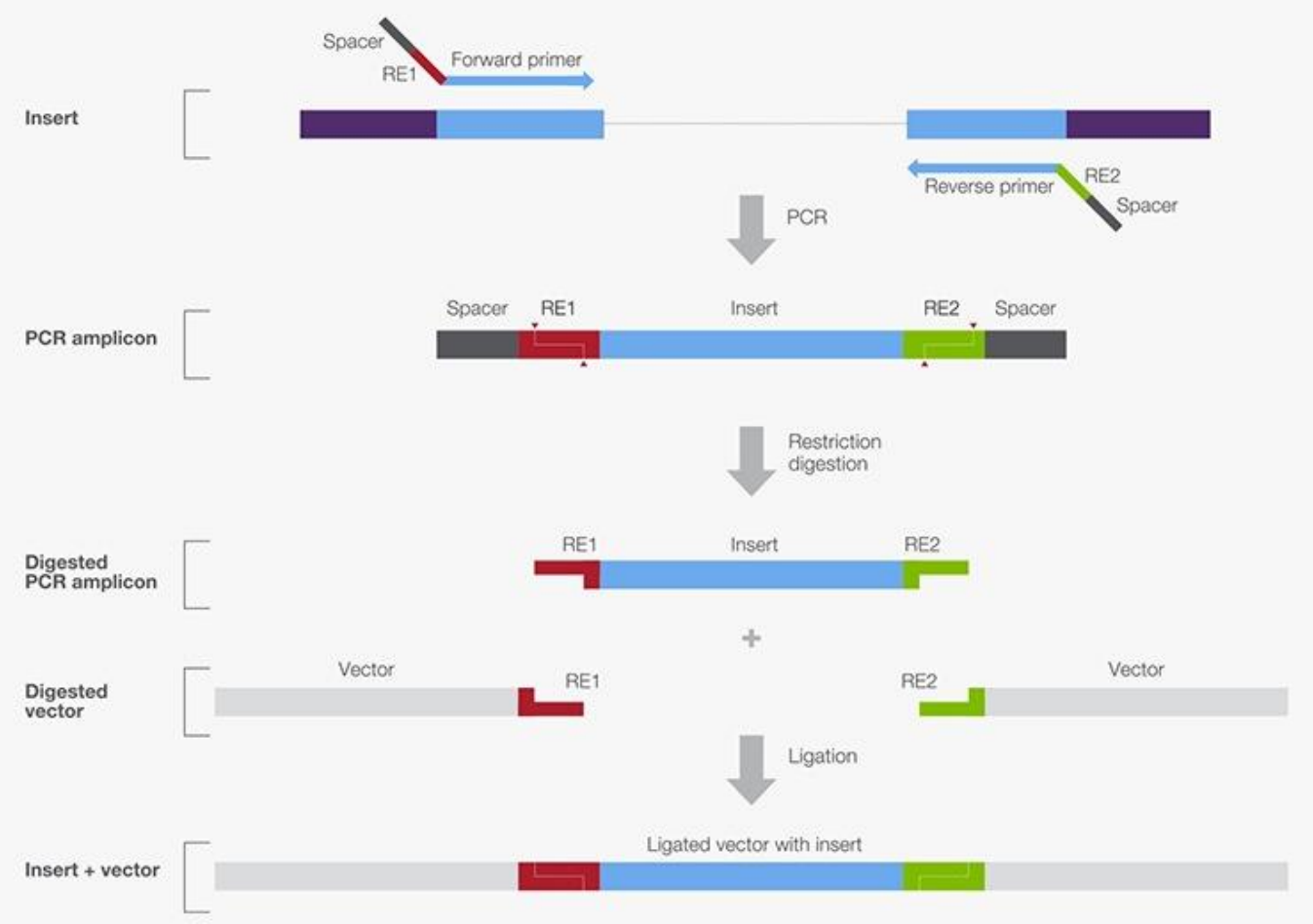
Price of synthetic fragments up to 3 kb is €0.10 per base pair.

Restriction cloning





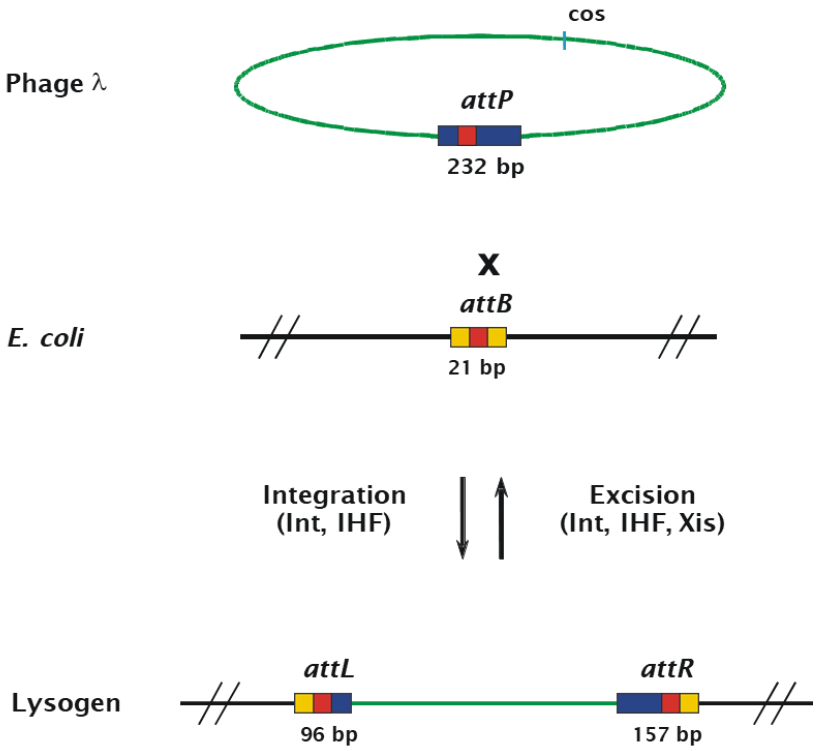
PCR-based cloning

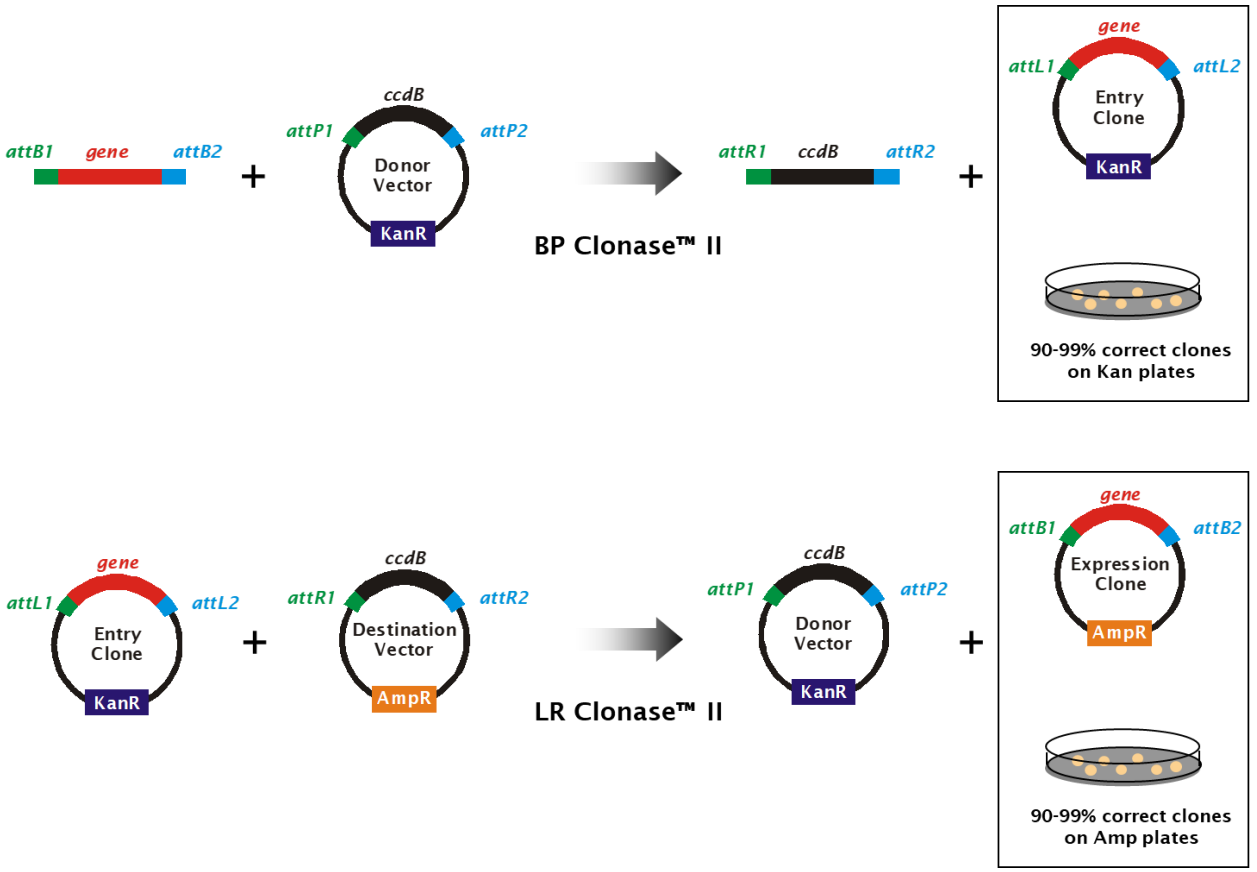


- DNA of interest can be flanked by any restriction site.

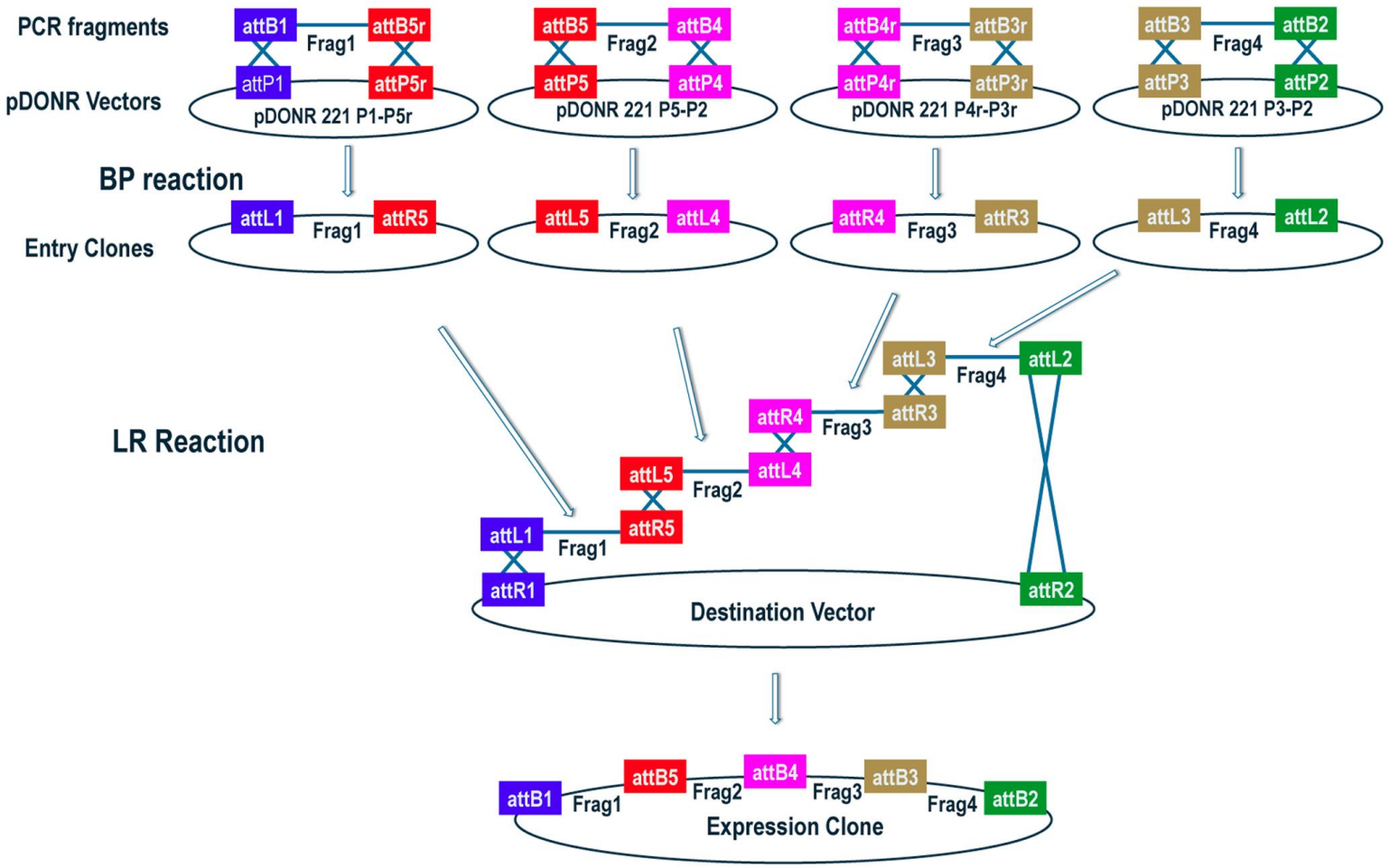
Gateway cloning

λ bacteriophage site specific integration system



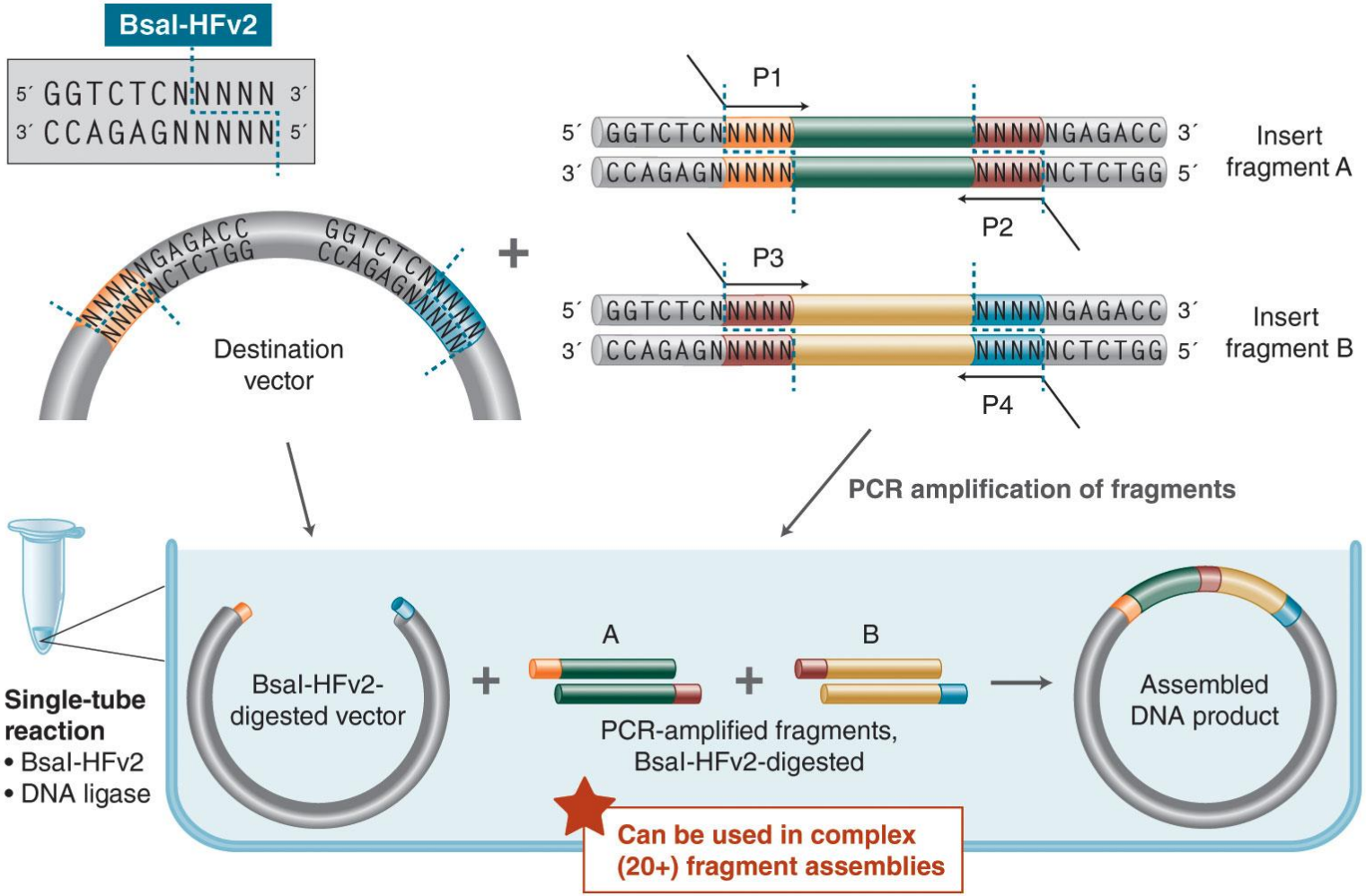


DNA assembly with multisite Gateway cloning



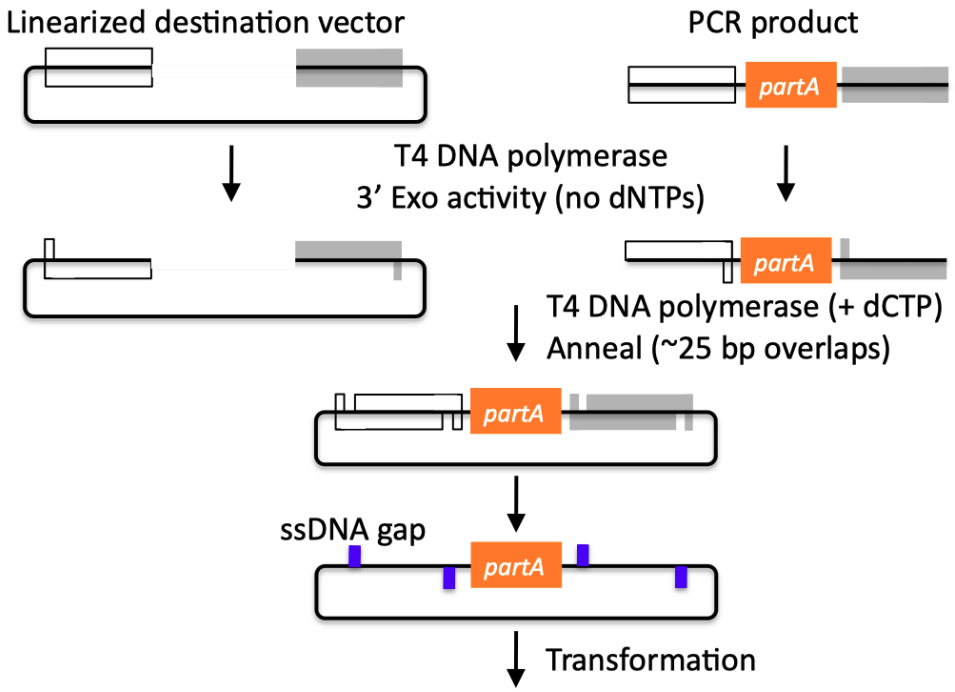
The multisite Gateway system relies on five sets of specific and non-crossreacting *att* sequnces.

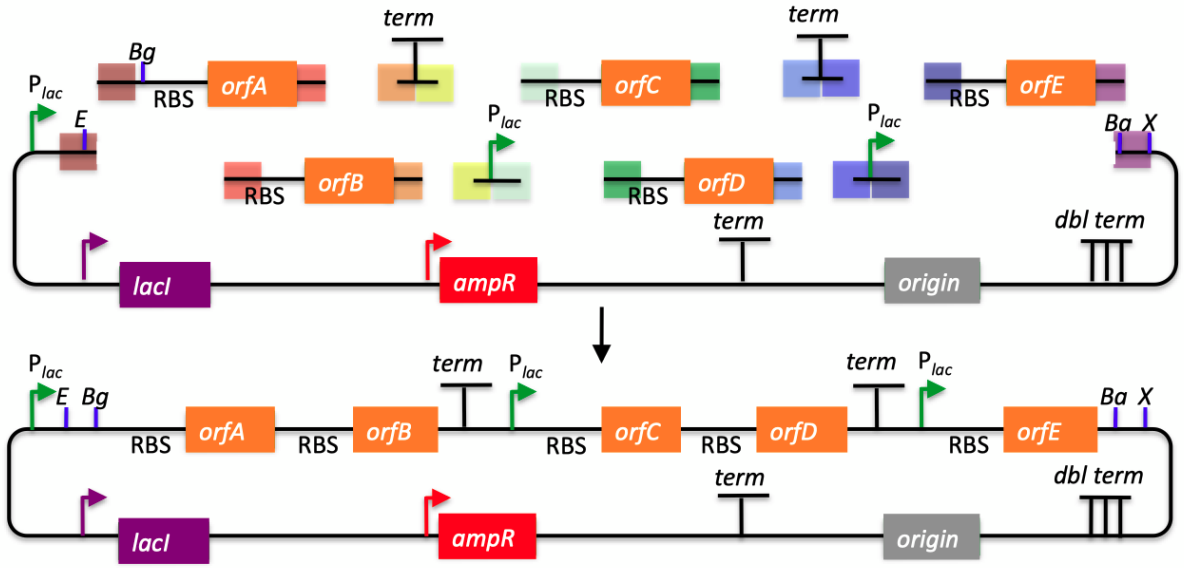
Golden Gate Assembly



- Seamless assembly

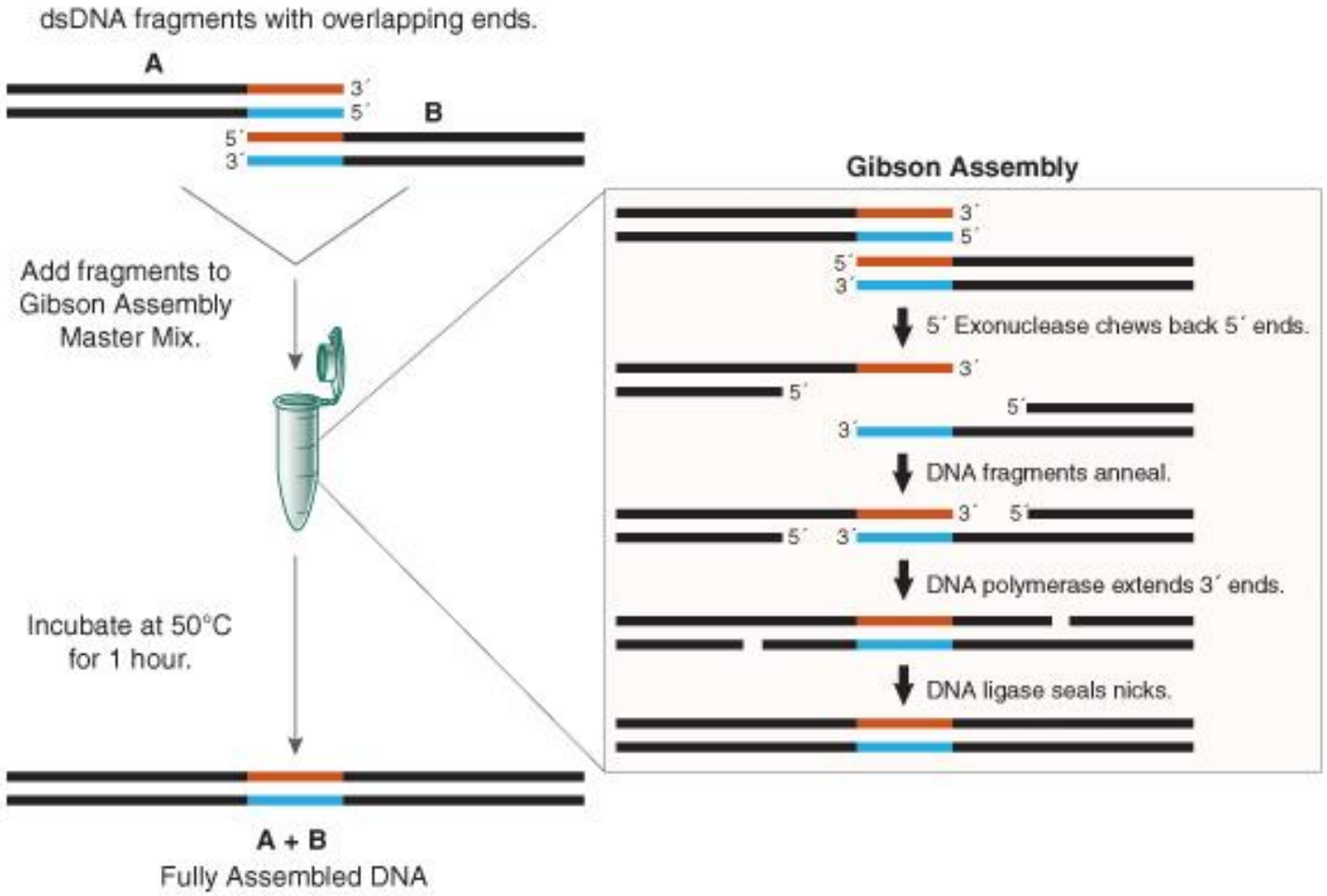
Sequence and ligation independent cloning (SLIC)





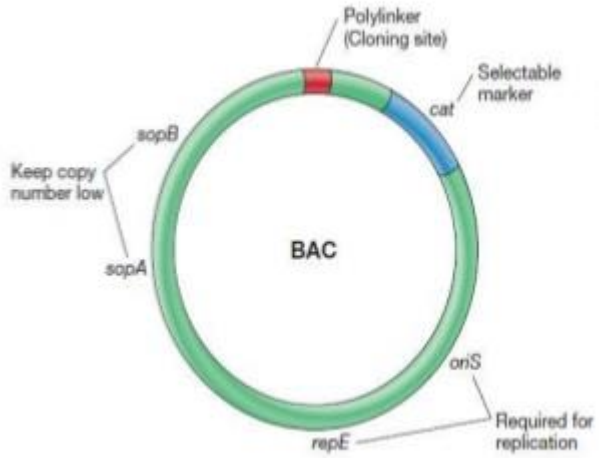
Gibson Assembly

- Seamless assembly



Gibson assembly was used to synthetise 16.3 kb mouse mitochondrial genome directly from 60-mer oligos.

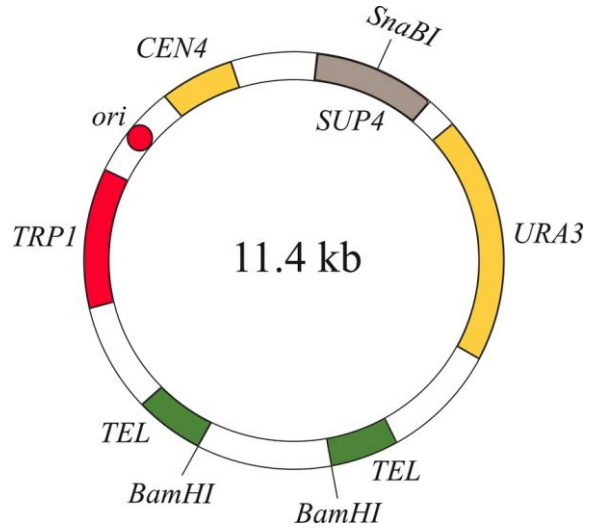
Bacterial artificial chromosomes (BACs)



- BACs are plasmids costructed with the replication origin of E. Coli F-factor, and so can be maintained in single copy per cell
- They can keep DNA fragments up to 300 kb
- Recombinant BACs are transformed into *E. coli* by electroporation
- Once in the cell, BAC replicates as an F-factor

Yeast artificial chromosomes (YACs)

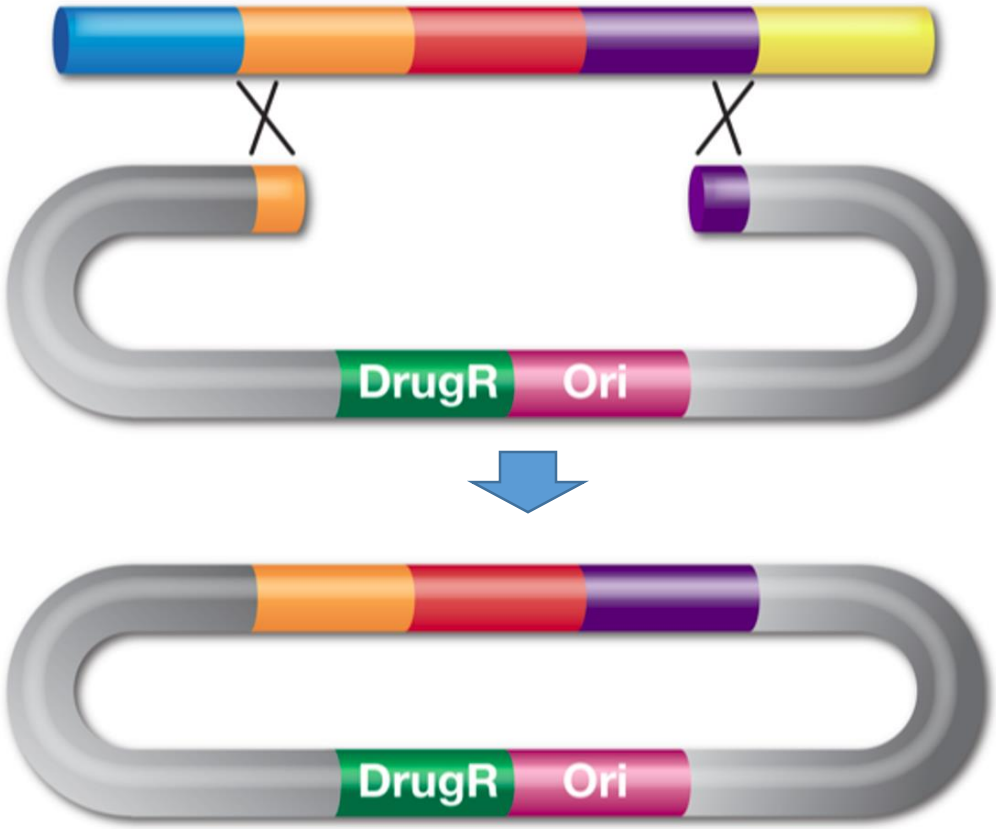
- YACs contain yeast centromeric DNA, telomeres and selctable markers
- YACs cab be shittled between yeast and *E. coli*
- They can keep DNA fragments up to 2000 kb



In vivo cloning: Recombineering

Recombineering takes advantage of cellular homologus recombination machinery to assemble desired DNA molecules in cells (bacteria, yeast) based on sequence homology.

In bacteria can efficincy of homologous recombination further be increased by coexpression of phage recombinat system (lambda red recombineering)

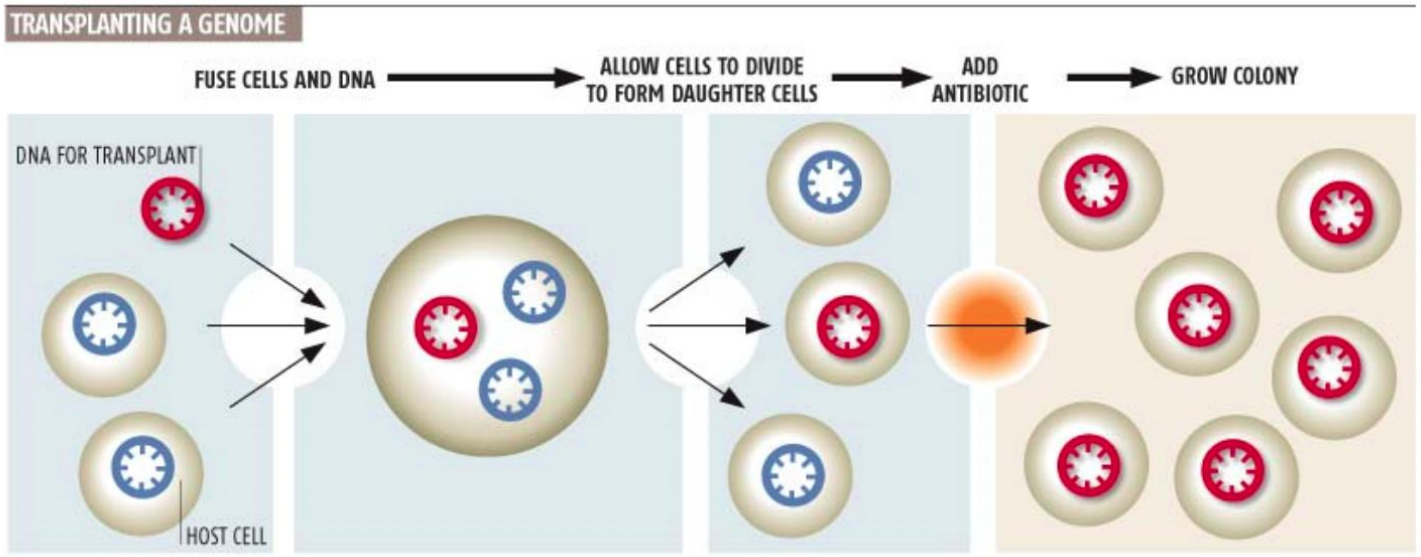


Towards synthetic genomes: genome transplantation

Mycoplasma is a genus of bacteria that lack a cell wall around their cell membranes. This characteristic makes them naturally resistant to antibiotics that target cell wall synthesis. Mycoplasmas have very small genomes.

Lartigue et al: Sceince 2007: replaced the genome of *Mycoplasma capricolum* with one from another species - *Mycoplasma mycoides* virtually transforming one spceies in another.

- Intact circular genome free of proteins was isoleted through puls field gel electrophoresis.
- Isolted genome of *M. mycoides* was transplanted into *M. capricolum* cells by polyethylene glycol–mediated transformation.



Selection: Tetracycline resistance
Screening: β-galactosidase (lacZ) genes
Main challenge: isolating intact genome and avoiding recombination

Assembly of a synthetic genome of *Mycoplasma genitalium*

- *Mycoplasma genitalium:* small pathogenic bacterium that lives on the skin cells of the urinary and genital tracts in humans. It has 600 kb genome encoding 500 genes- the smallest genome of a free living creature.
- Overlapping “cassettes” of 5 to 7 kilobases (kb), assembled from chemically synthesized oligonucleotides, were joined by in vitro recombination to produce intermediate assemblies of approximately 24 kb, 72 kb (“1/8 genome”) 144 kb (“1/4 genome”), which were all cloned as bacterial artificial chromosomes in Escherichia coli.
- Clones of all four 1/4 genomes was assembled by transformation associated recombination cloning in the yeast Saccharomyces cerevisiae, then isolated and sequenced.

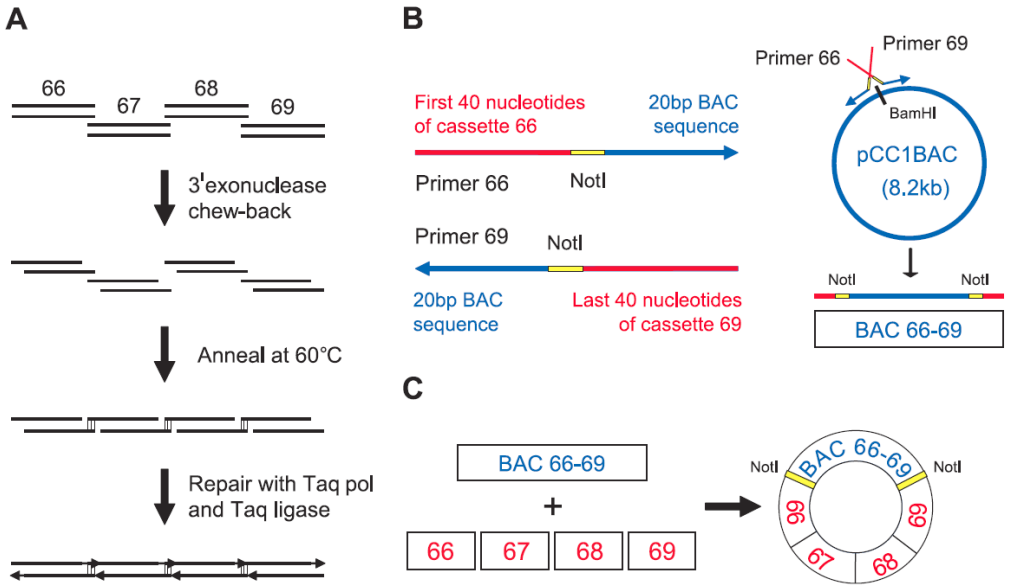
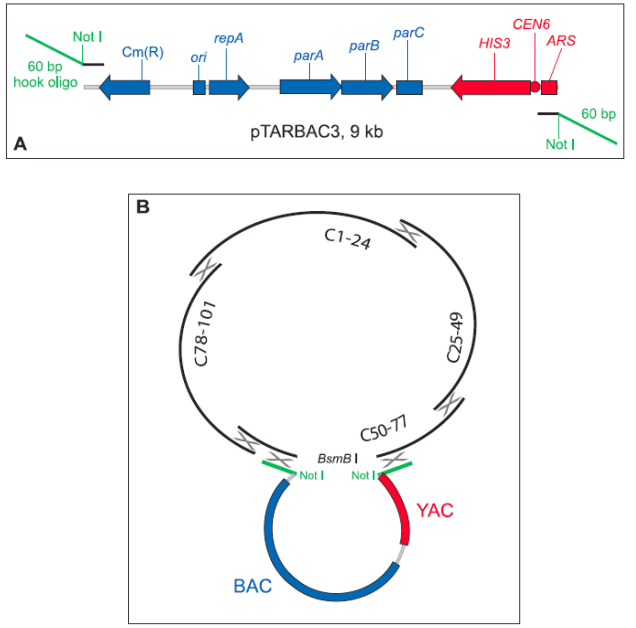


Fig. 3. Assembly of cassettes by in vitro recombination. (**A**) Diagram of steps in the in vitro recombination reaction, using the assembly of cassettes 66 to 69 as an example. (**B**) BAC vector is prepared for the assembly reaction by PCR amplification using primers as illustrated. The linear amplification product, after gel purification, is included in the assembly reaction of (A), such that the desired assembly is circular DNA containing the four cassettes and the BAC DNA as depicted in (**C**).



Whole genome assembly in yeast

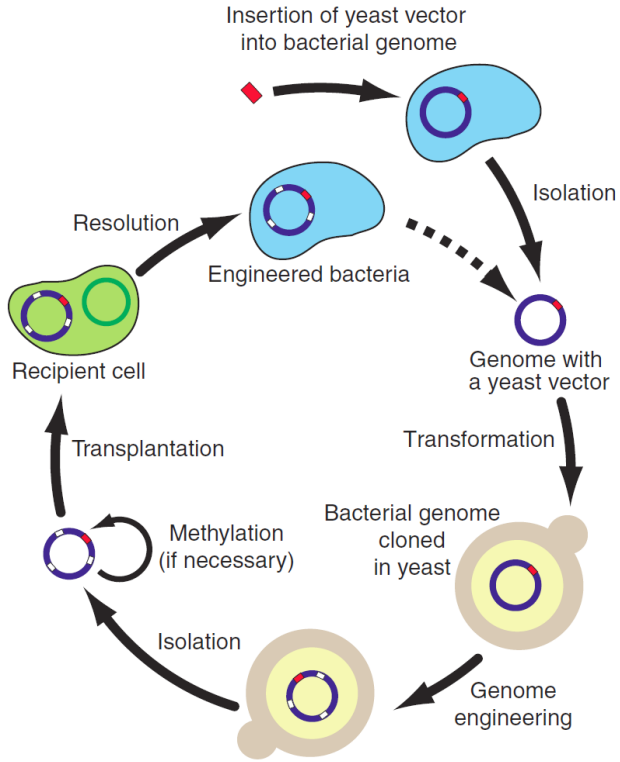
Gibson et al: Science 2008

Creation of bacteria with fully synthetic genome

- *M. genitalium* grows so slow that one experiment can take weeks to complete. The team decided to change microbes in midstream, sequencing the 1-million-base genome of the faster-growing *M. mycoides* and beginning to build a synthetic copy of its chromosome.

- After engeneering artificial genome in yeast, one must be able to transplant the bacterial chromosome from yeast back into a recipient bacterial cell.

Fig. 3. Moving a bacterial genome into yeast, engineering it, and installing it back into a bacterium by genome transplantation. A yeast vector is inserted into a bacterial genome by transformation. That genome is cloned into yeast. After cloning, the repertoire of yeast genetic methods is used to create insertions, deletions, rearrangements, or any combination of modifications in the bacterial genome. This engineered genome is then isolated and transplanted into a recipient cell to generate an engineered bacterium. Before transplantation it may be necessary to methylate the donor DNA in order to protect it from the recipient cell's restriction system(s). This cycle can be repeated starting from the newly engineered genome (dashed arrow).



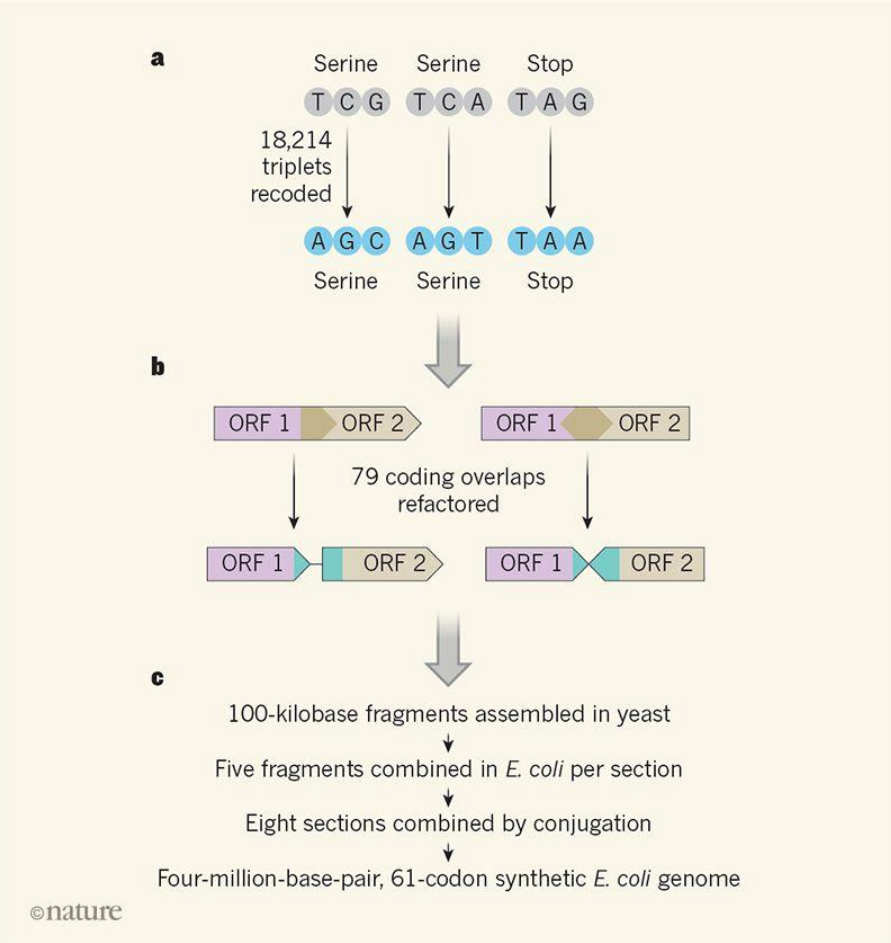
Lartigue et al: Science 2009

It took geneticist Craig Venter 15 years and US$40 million to synthesize the genome of a bacterial parasite. Venter's team wrote a few coded 'watermarks' into the genome sequence, which spelled out the names of the team members, as well as several famous quotes. But besides these tweaks and a few other changes, the synthetic M. mycoides genome was identical to its blueprint.

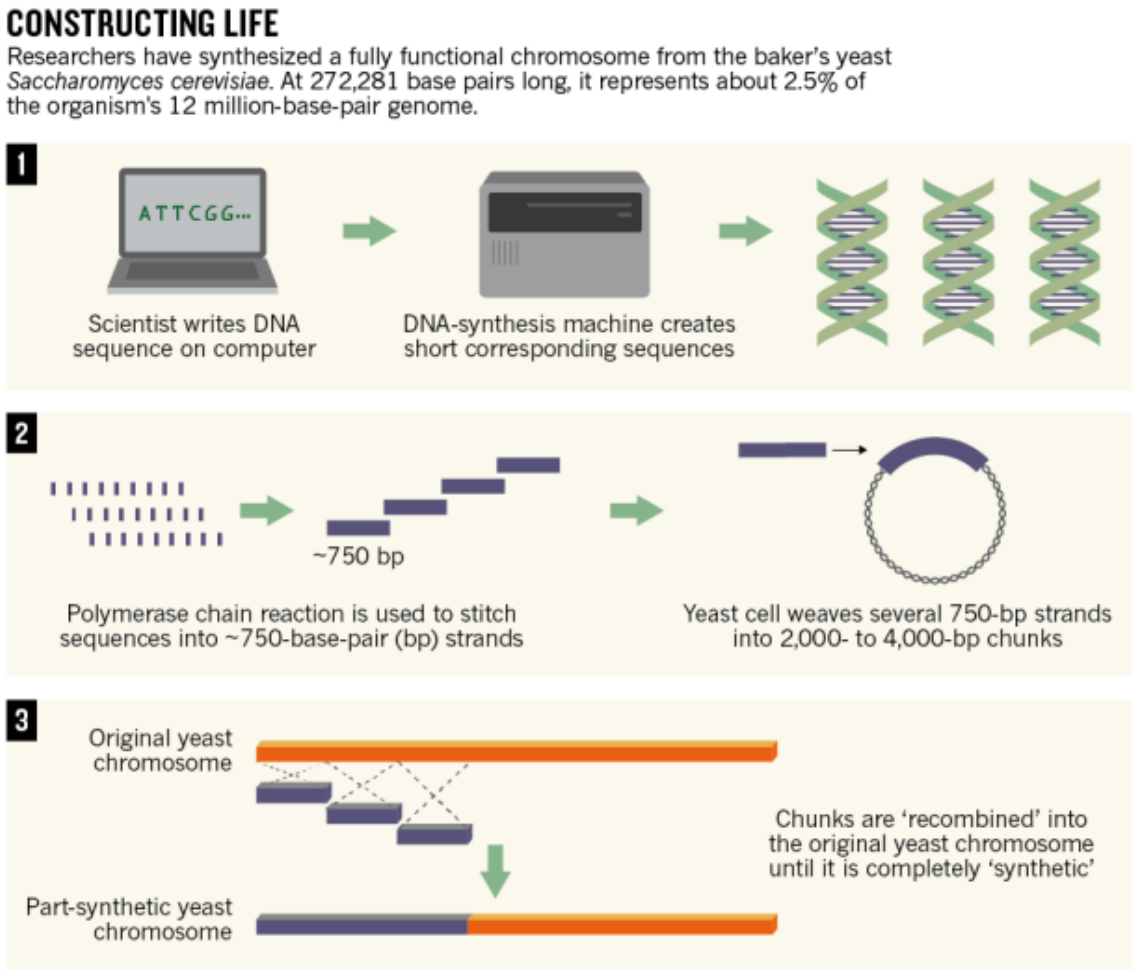
Synthetic genome of *E. coli* with reduced codon usage

- DNA was computationally designed, chemically synthesized and assembled in 100-kilobase fragments in vectors in *S. cerevisiae*
- These vectors were then taken up by E. coli and integrated into the genome in the direct place of the equivalent natural region. Iterating this process five times resulted in 500-kilobase sections of DNA being replaced by synthetic versions.
- Eight strains of *E. coli* were produced in this way, each harbouring synthetic DNA sections that covered a different region of the genome. These sections were then combined using conjugation to make the complete synthetic genome.
- Synthetic *E. coli* genome that uses only 61 of the 64 available codons in its protein-coding sequences, replacing two serine codons and one stop codon with synonyms.

Fredens et al: Nature 2019



An udergrad student project: Synthetic yeast chromosome



Annaluru et al., Science 2014

The 272,871–base pair designer eukaryotic chromosome, synIII, which is based on the 316,617–base pair native *S. cerevisiae* chromosome III.

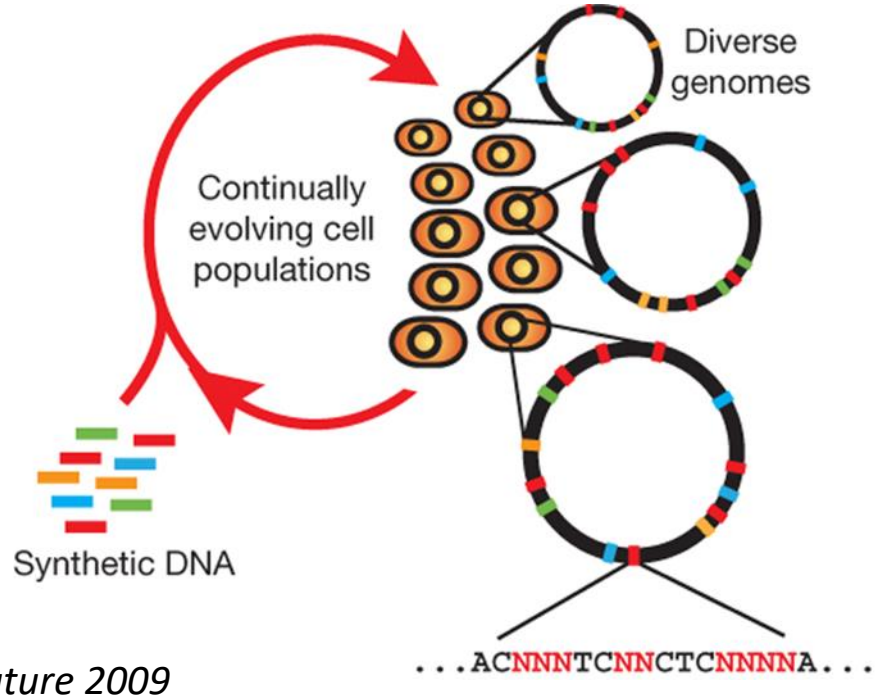
Each student makes their own stretch of the yeast genome, which involves stitching together very short lengths of DNA created by a DNA-synthesis machine into ever-larger chunks. These chunks are then incorporated into the yeast chromosome, a few at a time, through a process called homologous recombination. Eventually, this results in an entirely synthetic chromosome.

Genome editing technologies

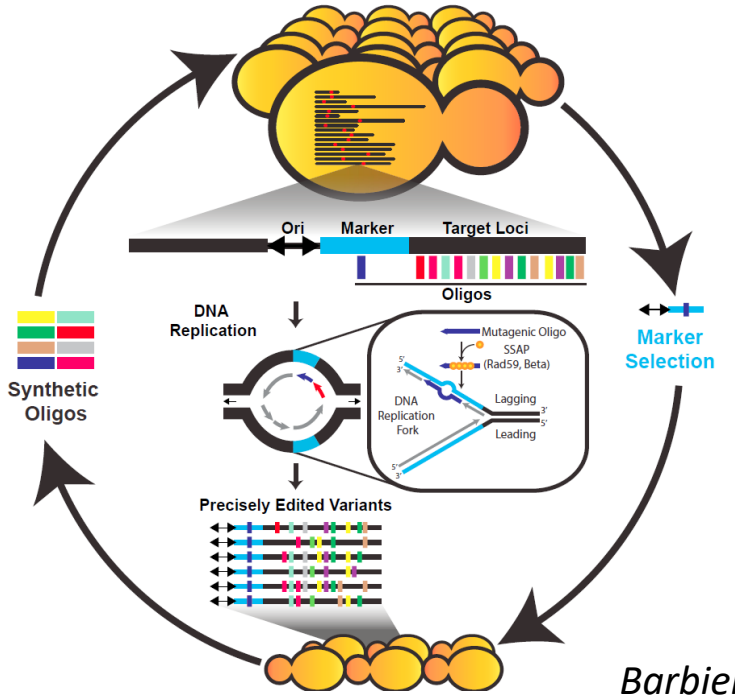
- Synthetic genome technology is still expensive, laborious and limited to a few organims.
- Genome editing is affordable, relatively easy and applicable in a wide range of organisms.

MAGE: Multiple automated genomic engineering

- Allelic conversion mediated by directing oligonucleotides to the lagging strand of the replication fork during DNA replication that are delivered to a cell by electroporatin.
- The conversion is facilitated by the bacteriophage λ-Red ssDNA-binding protein β in *E. coli* and conversion effcinecies can reach up to 30%.
- Up to 50 loci can be simultaneously modified in a single cell.
- MAGE enables the rapid and continuous generation of sequence diversity at many targeted chromosomal locations across a large population of cells through the repeated introduction of synthetic DNA.
- MAGE has also beed developed for budding yeast with symultaneous incorporation of up to 12 oligos

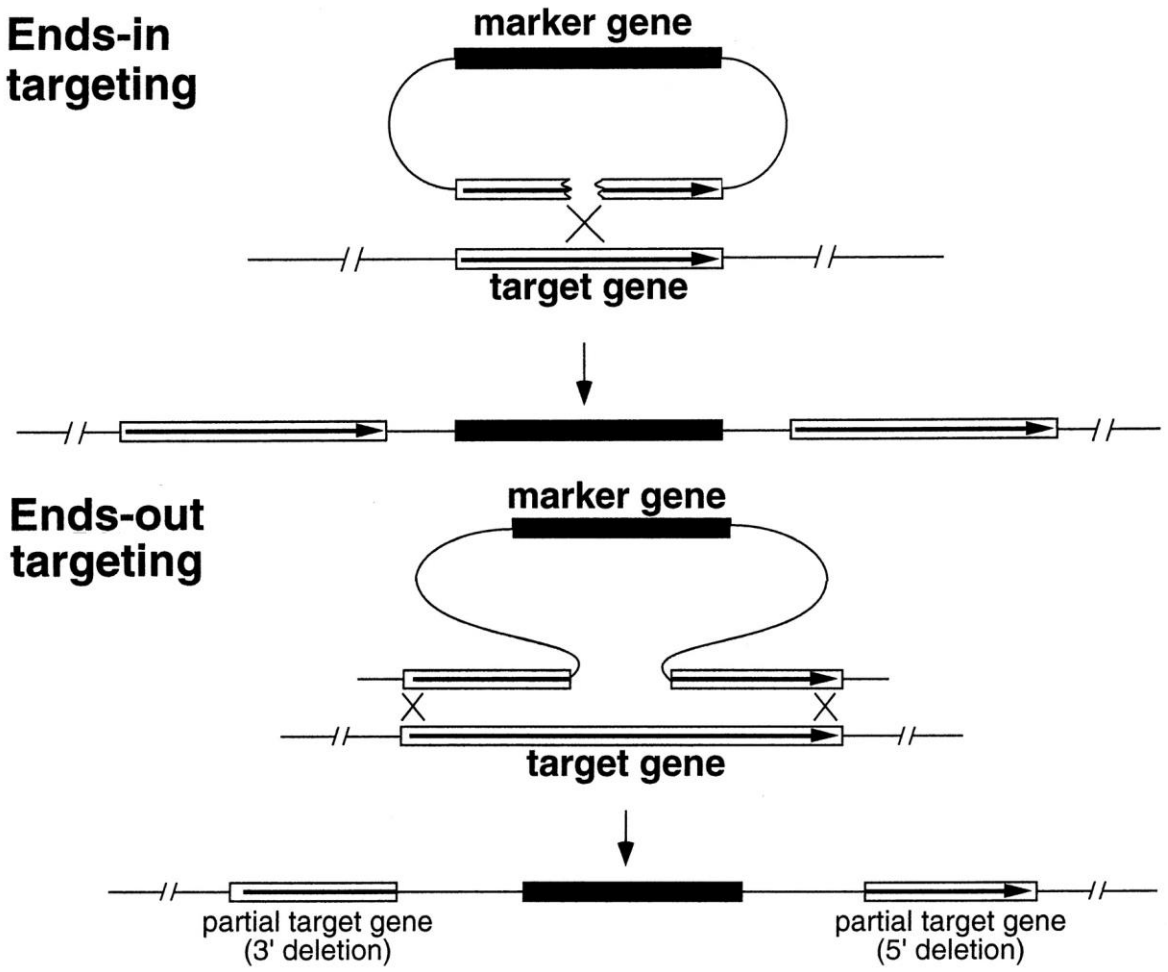


Wang et al., Nature 2009



Barbieri et al., Cell 2017

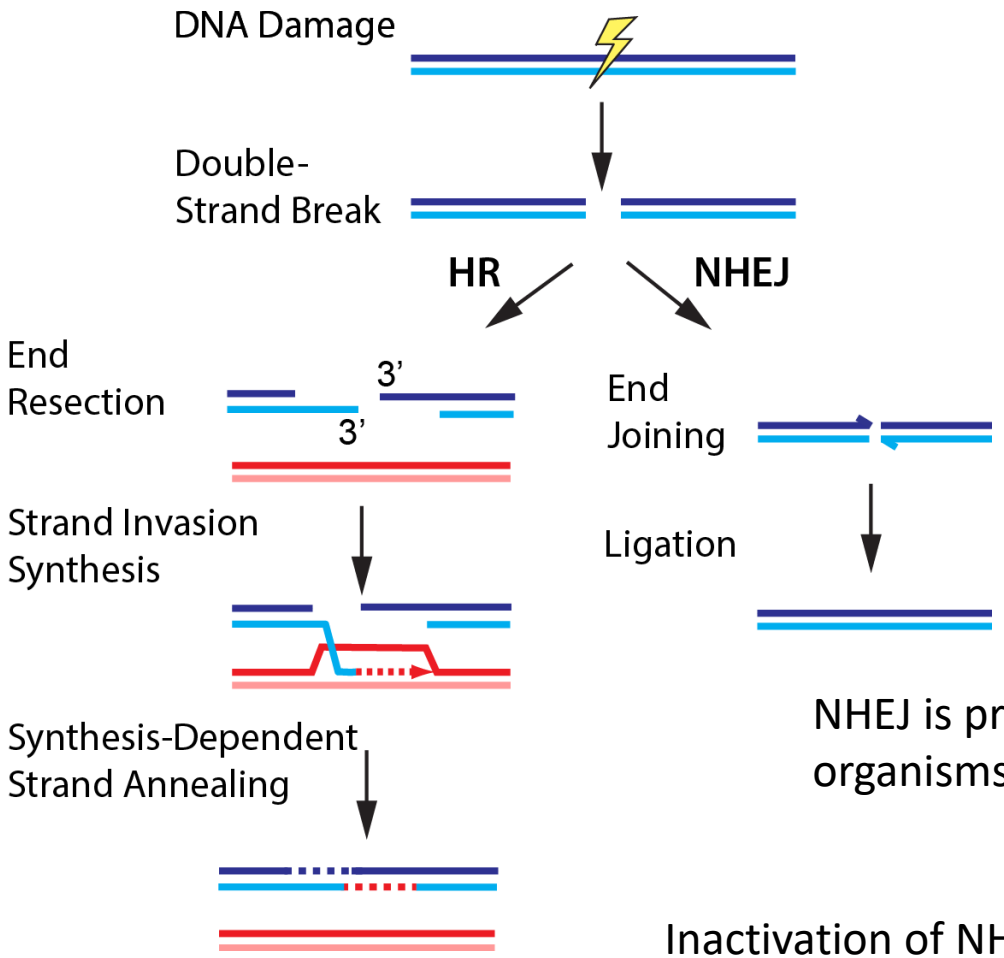
Gene targeting



- Targeted gene knock-outs
- Conditional gene knoc-outs
- Gene tagging
- Amino acid substitutions
- ...

Gene targeting is extremely ineffcient in majority of eukaryotes

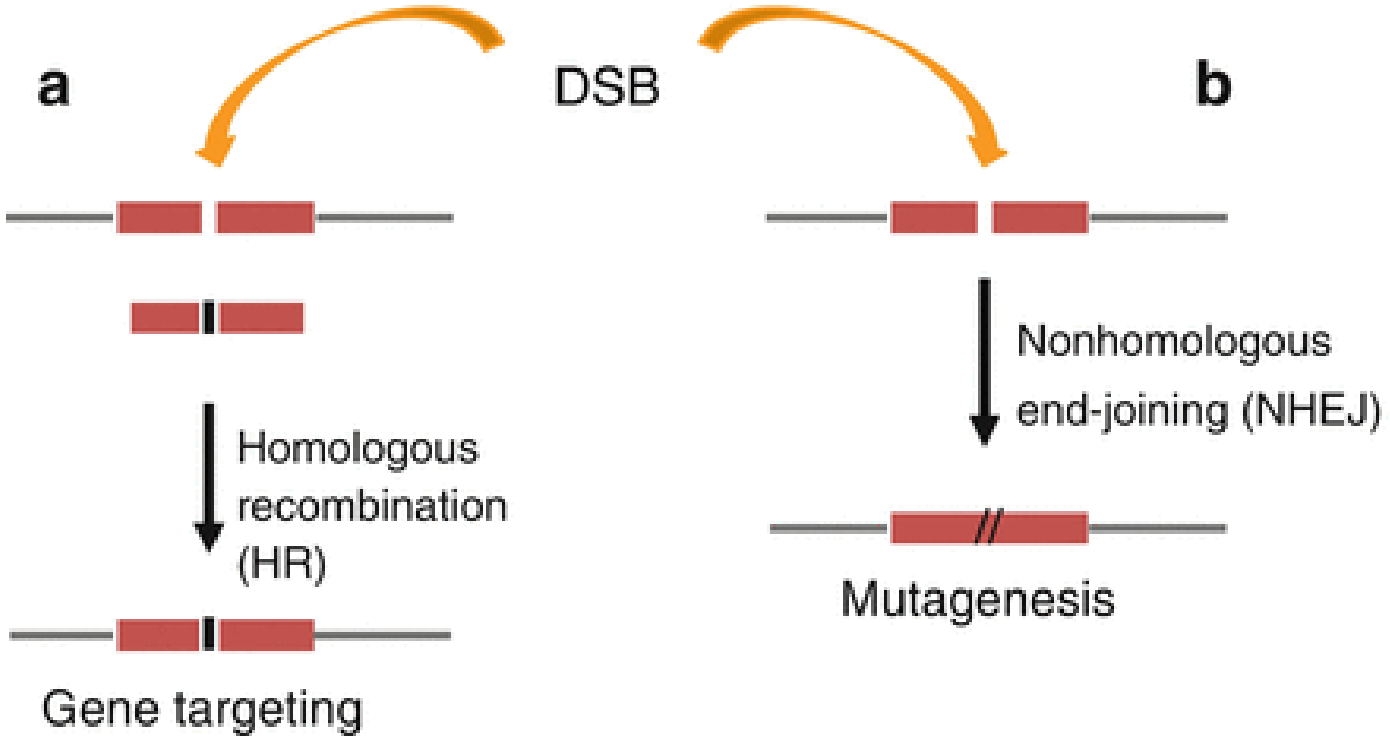
Efficient gene targeting is limited to a few model systems (yeast, ES in mouse, chicken DT40 cell line, moss *Physcomitrella patens, ...*)



NHEJ is prevalent mode of repair in organisms with complex genome

Inactivation of NHEJ can boost GT in some organisms

Break induction in targeted locus boosts gene targeting

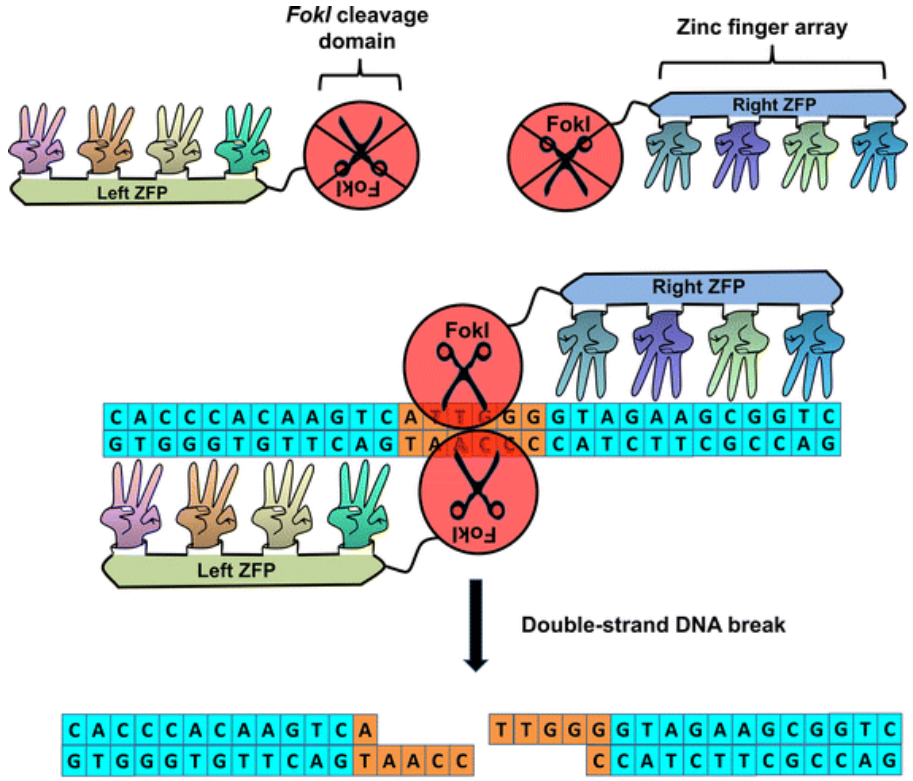


➡ Interest in programmable nucleases that can cleave sequence of choice.

Zinc-finger nucleases

Zinc-fingr: a small protein strucutral motif coordinated by one or more zinc ions. Involved in DNA/RNA binding. A singel zinc finger can recognize 2 or 3 bases, but does not bind very tightly to nucleic acids. DNA binding specificity is achived by combining multiple zinc-fingers.

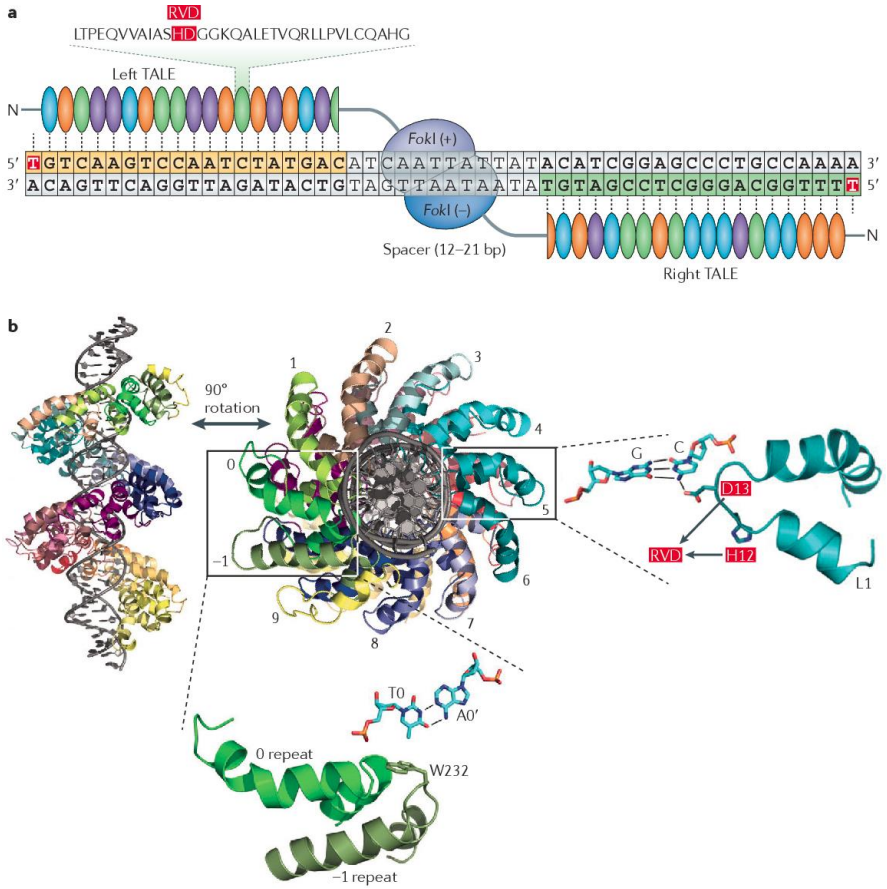
Zinc-finger nucelases are formed by combining multiple zinc-fingers with the nuclease domain of FokI restriction endonuclease. FokI acts as a dimer.



- Expensive, laborious and time consuming in making construts
- Target each site separately
- Low efficiency
- Fewer off-target mutations
- Target any genomic location

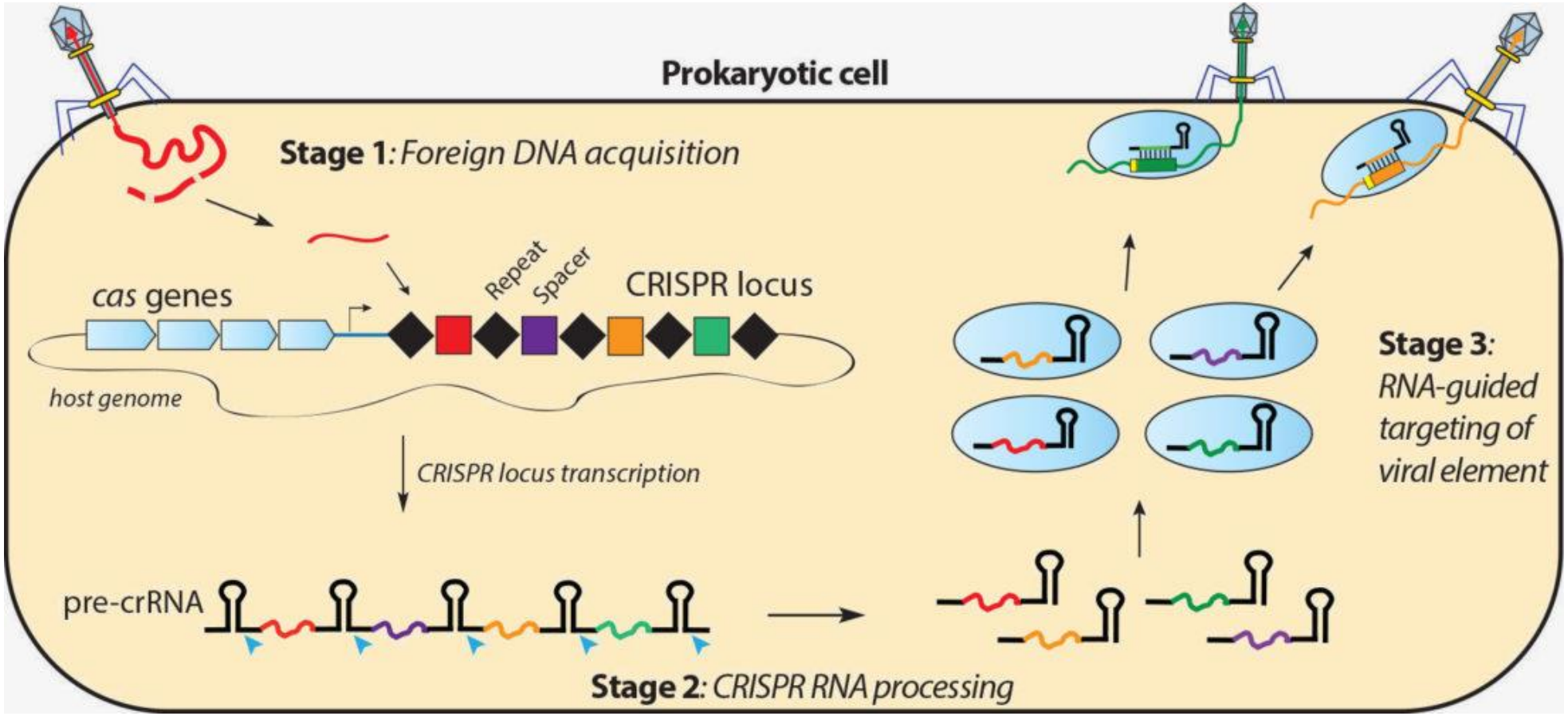
Transcription activator-like effector nucleases (TAELNs)

TAL-effectors: proteins are secreted by *Xanthomonas* bacteria when they infect plants, where they bind to promoters and activate genes that aid infection. The DNA binding domain contains a repeated highly conserved 33–34 amino acid sequence with divergent 12th and 13th amino acids. These two positions, referred to as the Repeat Variable Diresidue (RVD), are highly variable and show a strong correlation with specific nucleotide recognition.

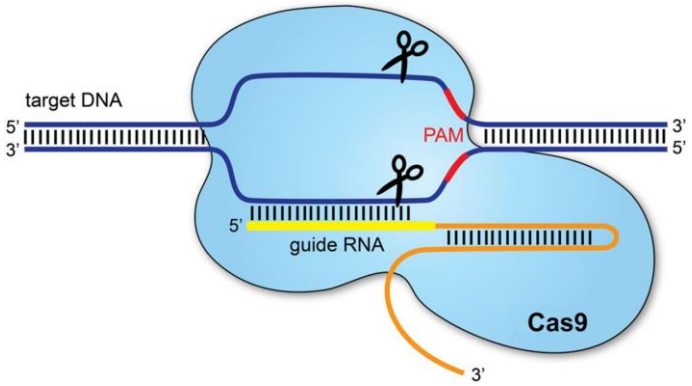


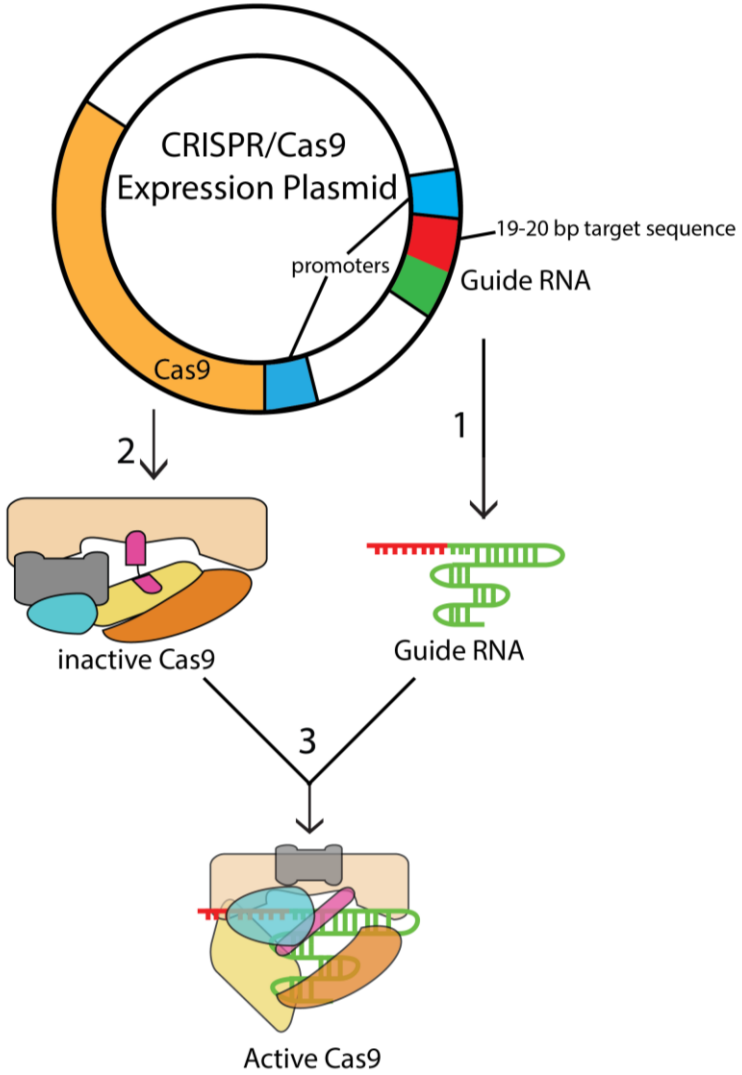
- Expensive, laborious and time consuming in making construts (easier than zinc-fingers)
- Target each site separately
- Low efficiency
- Sensitive to DNA methylation
- Fewer off-target mutations
- Target any genomic location

CRISPR-Cas: Prokaryotic adaptive immune system



Programmable CRISPR-Cas9 nuclease

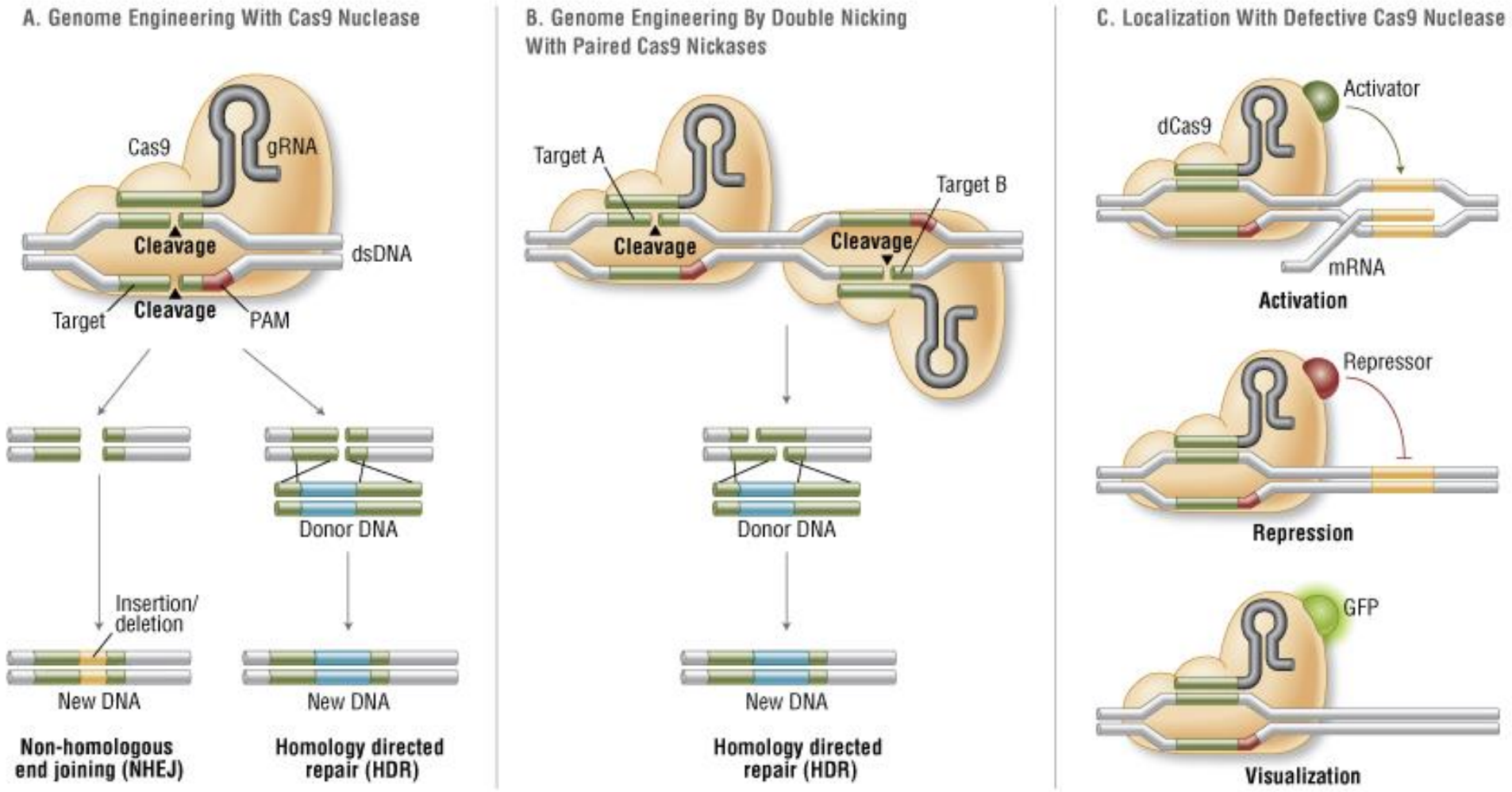




- Target site must have PAM sequence
- Off target cleavage
- Cheap and easy to make costructs
- Highly versitile

It works!!!!

Applications of CRISPR-Cas9 technology



Multiplex genome engineering using CRISPR/Cas systen

Simultaneous expression of multiple sgRNAs allows effcient targeting of entire gene families.

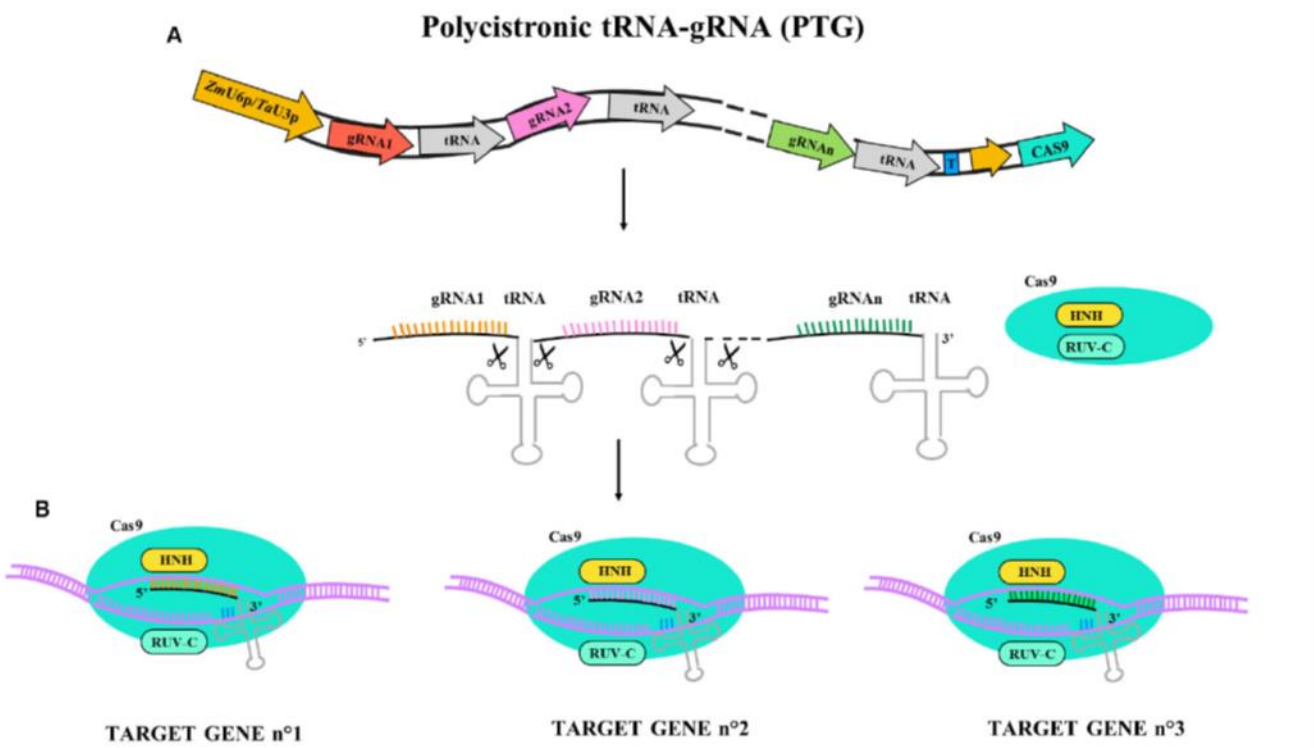
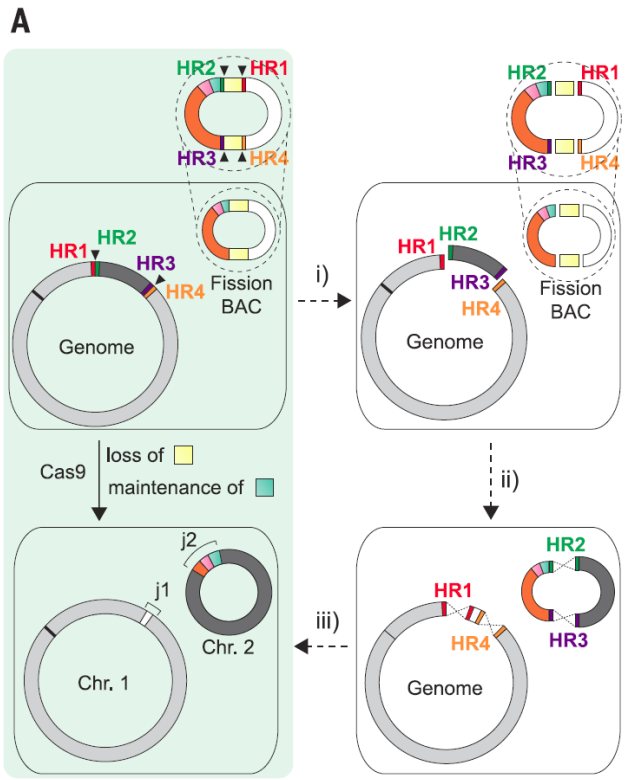


FIGURE 2 | Illustrative diagram of polycistronic tRNA-gRNA (PTG) gene construct and targeting activity for Cas9. PTG is composed of t-RNA-gRNA repeats and is upregulated by ZmU6 promoter or TaU3 promoter according the experimental design as different terminator region (T) are adopted. **(A)** PTG primary transcript. Endogenous endonuclease cuts the tRNA ends and let each tRNA-gRNA targeting the corresponding gene sequence. **(B)** In PTG system more sequence targets are available (n° gene targets) and the different gRNA are represented in different colors (orange, pink, and green).

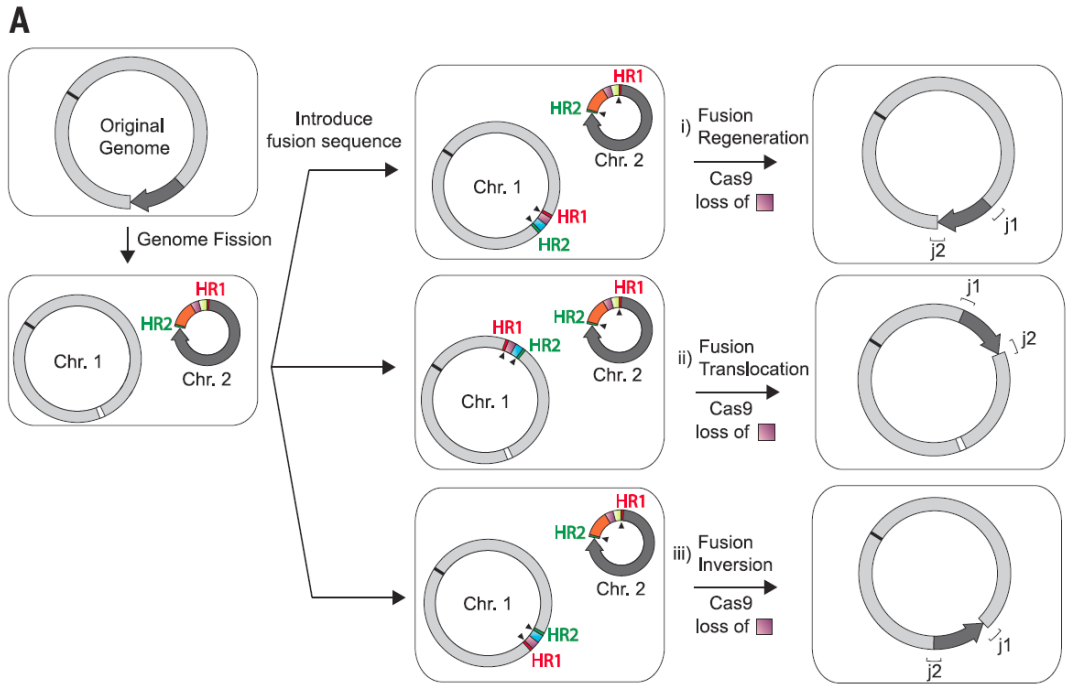
Genome engineering with the CRISPR-Cas system

Chromosome fission



(A) E. coli harbors a fission BAC containing a double selection cassette. During fission, (i) Cas9 induces six cuts (black triangles), splitting the genome into fragment 1 (light gray, containing oriC indicated by black line) and fragment 2 (dark gray) and the fission BAC into four pieces (linker sequence 1, linker sequence 2, and two copies of rpsL). (ii) Homology regions (HRs) between fragments and their cognate linkers. (iii) Lambda red recombination joins fragments and linkers to yield chromosomes 1 and 2 (Chr. 1 and Chr. 2). Junctions 1 and 2 (j1 and j2) are new junctions.

Chromosome fusion

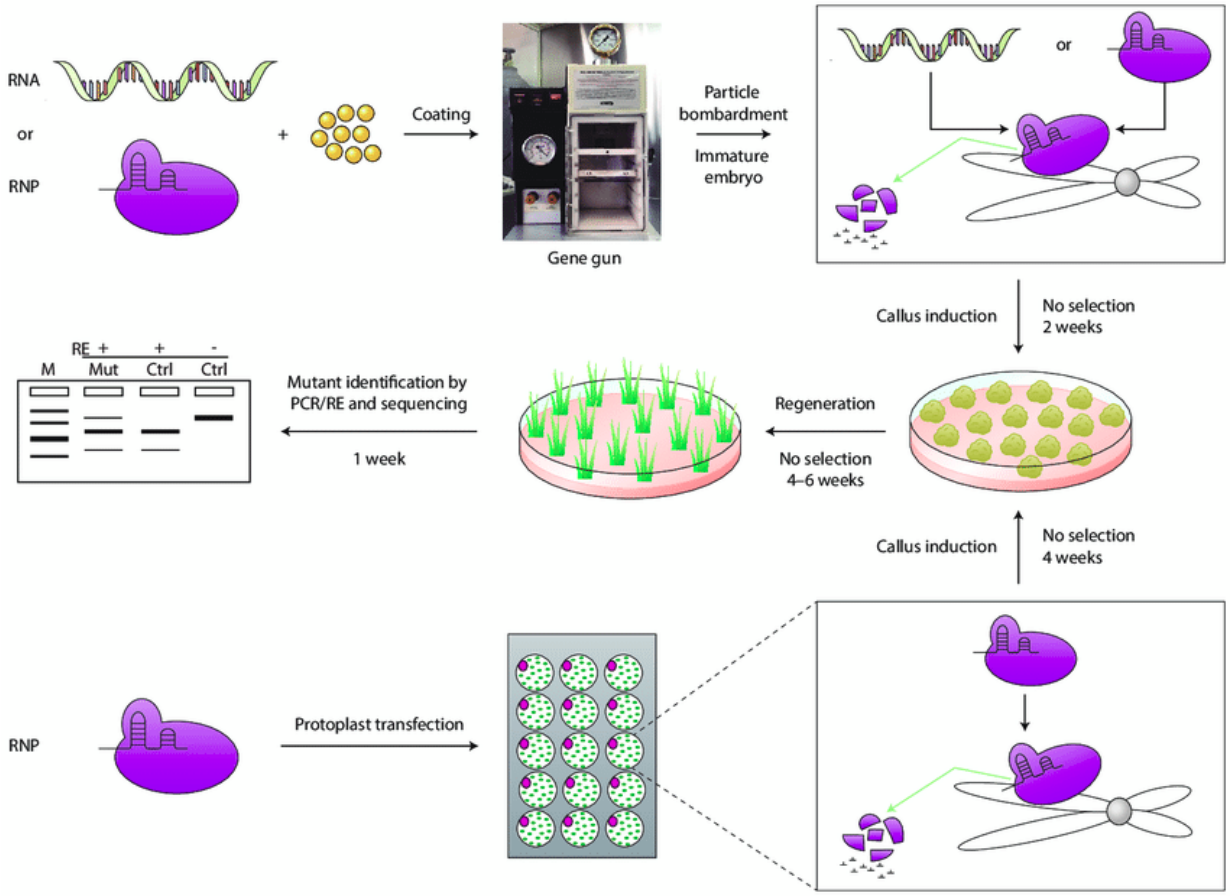


(A) E. coli with two chromosomes (Chr. 1 ~3.45 Mb and Chr. 2 ~0.54 Mb) was generated by fission. The sequence of Chr. 2 is watermarked as described in the text. The color-coding is as in Fig. 1A; a pheS*-KanR double selection cassette (purple and yellow, respectively) is shown. A fusion sequence, consisting of a pheS*-HygR (purple and blue, respectively) double selection cassette flanked by HR1 and HR2, is introduced in the indicated positions and orientation in Chr. 1 by lambda-red recombination. Cas9 spacer-directed cleavage (black arrows), lambda-red recombination, and selection for fusion products through the loss of pheS* on 4-chloro-phenylalanine yield the indicated products. (i) Regenerating the original genomic arrangement, (ii) translocation of the 0.54-Mb segment 700 kb away from its original position, and (iii) inversion of the 0.54-Mb segment.

Wang et al., Science 2019

Are genome edited organisms GMOs?

transgene free CRISPR/Cas editing



March 2019: A food service company in the Midwestern region of the United States is now using an oil made from **genetically edited soybeans** in its sauces, dressings, and fryer. By using gene-editing technology to deactivate two genes found in soybeans, Minnesota-based agriculture company Calyxt says it's created a soybean oil with no trans fats and more heart-healthy fats than traditional soybean oils.

Recommended reading:

Gibson et al. (2008) Complete Chemical Synthesis, Assembly, and Cloning of a Mycoplasma genitalium Genome. Science 319:1215

Lartigue et al. (2009) Creating Bacterial Strains from Genomes That Have Been Cloned and Engineered in Yeast. Science 325:1693

Marraffini et al. (2015) CRISPR-Cas immunity in prokaryotes. Nature 526:55