PLNT2530 (2024)
Unit 9

Genome Editing

Genome Editing

Definition: Genome Editing is the directed mutation of specific sequences in the genome.

Genome Editing is distinct from transformation, because nothing is added. A short chromosomal sequence is modified *in-situ*.

Genome Editing - Technologies

- **ZFN - Zinc finger nucleases** - Zinc Fingers are a class of DNA-binding proteins that recognize specific short sequences. They can be converted into sequence-specific nucleases (a bit like restriction endonucleases) by mutating finger domains to change sequence specificity, and joining the ZF coding domain with the nuclease domain of the restriction enzyme FokI.
- **TALENS** - TALE (transcription activator-like effectors) genes from bacteria have DNA binding domains that are very easy to re-engineer to recognize any short DNA sequence. Engineered TALE genes can be joined to enzymes like FokI nuclease to create sequence-specific nucleases.
- **CRISPR (clustered, regularly interspaced, short palindromic repeats)** - The sequence specificity of bacterial Cas complexes is determined by a short RNA called a crRNA, carried by the Cas9 protein. Direct synthesis of crRNAs makes it possible to target Cas9 to almost any DNA sequence. Thus, it is much easier and faster to generate new CRISPR specificities than using ZFN or TALENS.

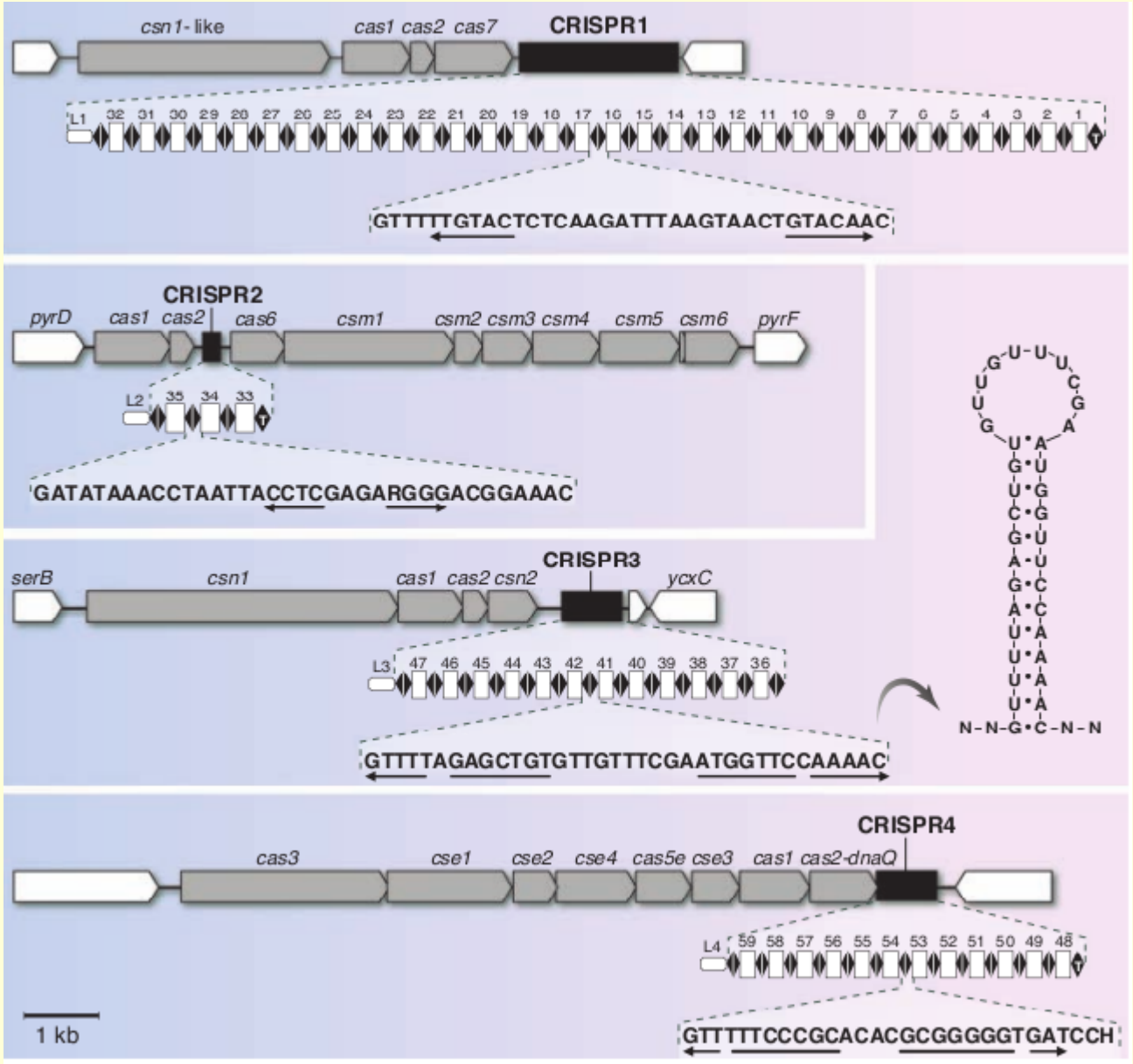
VIDEO: Genome Editing with CRISPR? https://youtu.be/2pp17E4E-O8
McGovern Institute for Brain Research at MIT

CRISPR loci in Bacteria

In many bacterial and archaeal genomes, loci have been identified containing a series of genes called cas genes, followed by arrays of short repeat-spacer-repeat motifs.

CRISPR arrays consist of conserved repeats, separated by unique sequences referred to as spacers. Each spacer sequence is derived from a different foreign sequence.

Right: four CRISPR arrays in *Streptococcus thermophilus*.



◆ - conserved repeats □ - spacer DNA of foreign origin

Horvath P and Barrangou R (2010) Science 327:167-170.

CRISPR loci in Bacteria

- repeats
  - 23 - 47 bp
  - typically inverted repeats that can form stem-loop structures
- spacers
  - 21 - 72 bp
  - derived from foreign sequences
- repeat-spacer units - typically about 50 copies per genome, but the largest known is from *Chloroflexus* Y-400-fl, at 375 units
- many microbial genomes contain several CRISPR loci
- CRISPR loci are typically found on the chromosome, but can be found on plasmids

CRISPRs were discovered to be a mechanism for adaptive immunity to viruses and other foreign sources of DNA or RNA

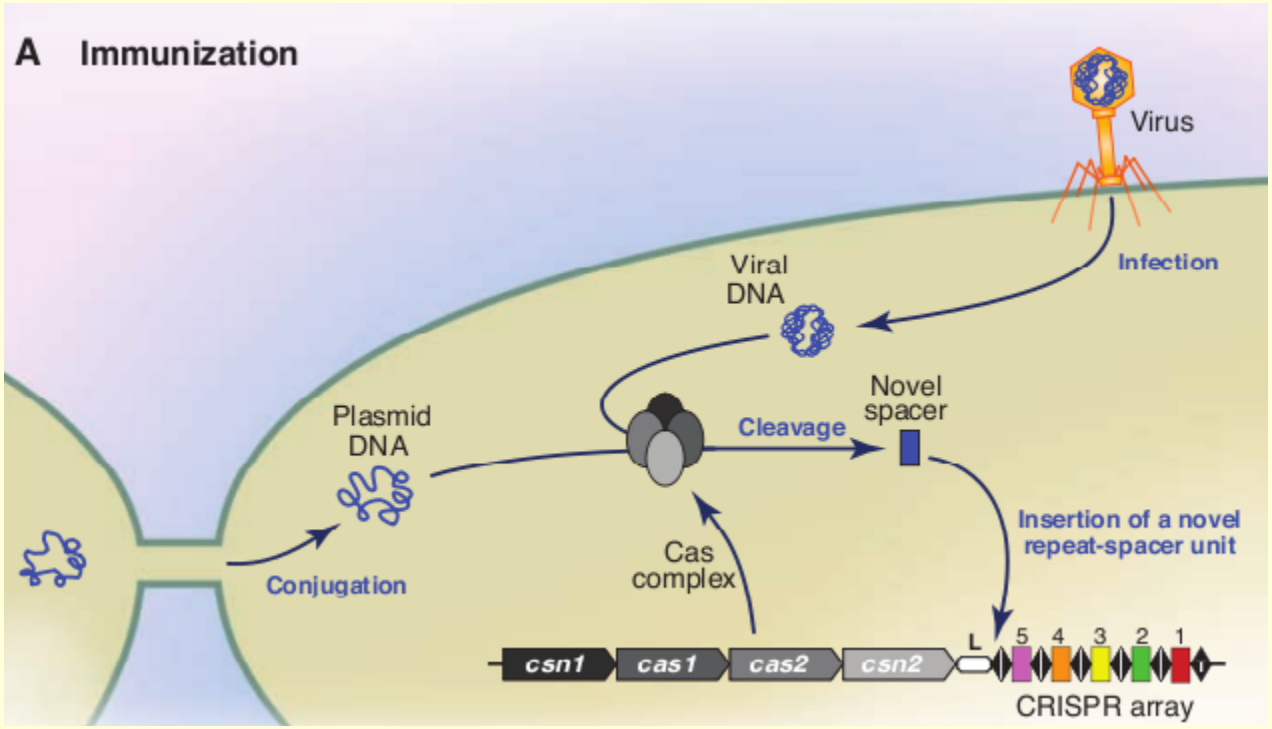
Horvath P and Barrangou R (2010) Science 327:167-170.

Aquisition of new CRISPR elements in Bacteria

DNA from infecting virus or foreign plasmid is cleaved in by Cas complex proteins.

Cleavage products are inserted as spacer DNA, 5' to the first repeat unit.

In this way, between 21 and 72 bases of foreign DNA are incorparated as part of the CRISPR array.



- conserved repeats
- spacer DNA of foreign origin

Horvath P and Barrangou R (2010) Science 327:167-170.

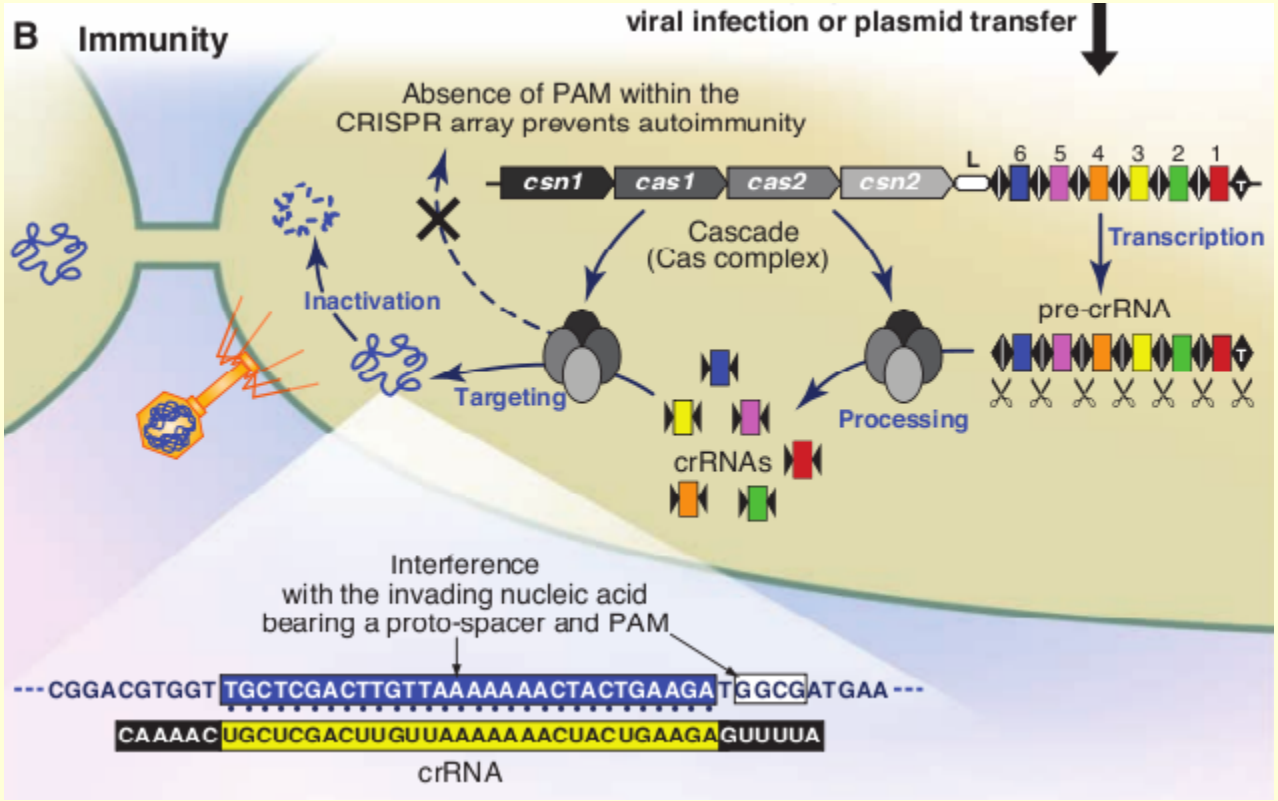
CRISPR-based immunity in Bacteria

CRISPR loci are transcribed as a single pro-crRNA, and the RNA is cleaved by cas proteins into short crRNAs.

Each crRNA contains a single CRISPR spacer, derived from prior contact with foreign DNA such as viruses.

A cas complex binds the crRNA.

If the crRNA on the cas complex can base-pair with foreign DNA, the DNA is cleaved by cas9.



black diamonds - conserved repeats
coloured boxes - spacer DNA of foreign origin

Horvath P and Barrangou R (2010) Science 327:167-170.

Genome Editing - Simplified Cas9-sgRNA editing system

The CRISPR locus from *Streptococcus pyogenes* has been adapted to create a minimal genome editing system.

The only protein needed is Cas9. Like most proteins that modify DNA, Cas9 is huge: MW=158,441 daltons.

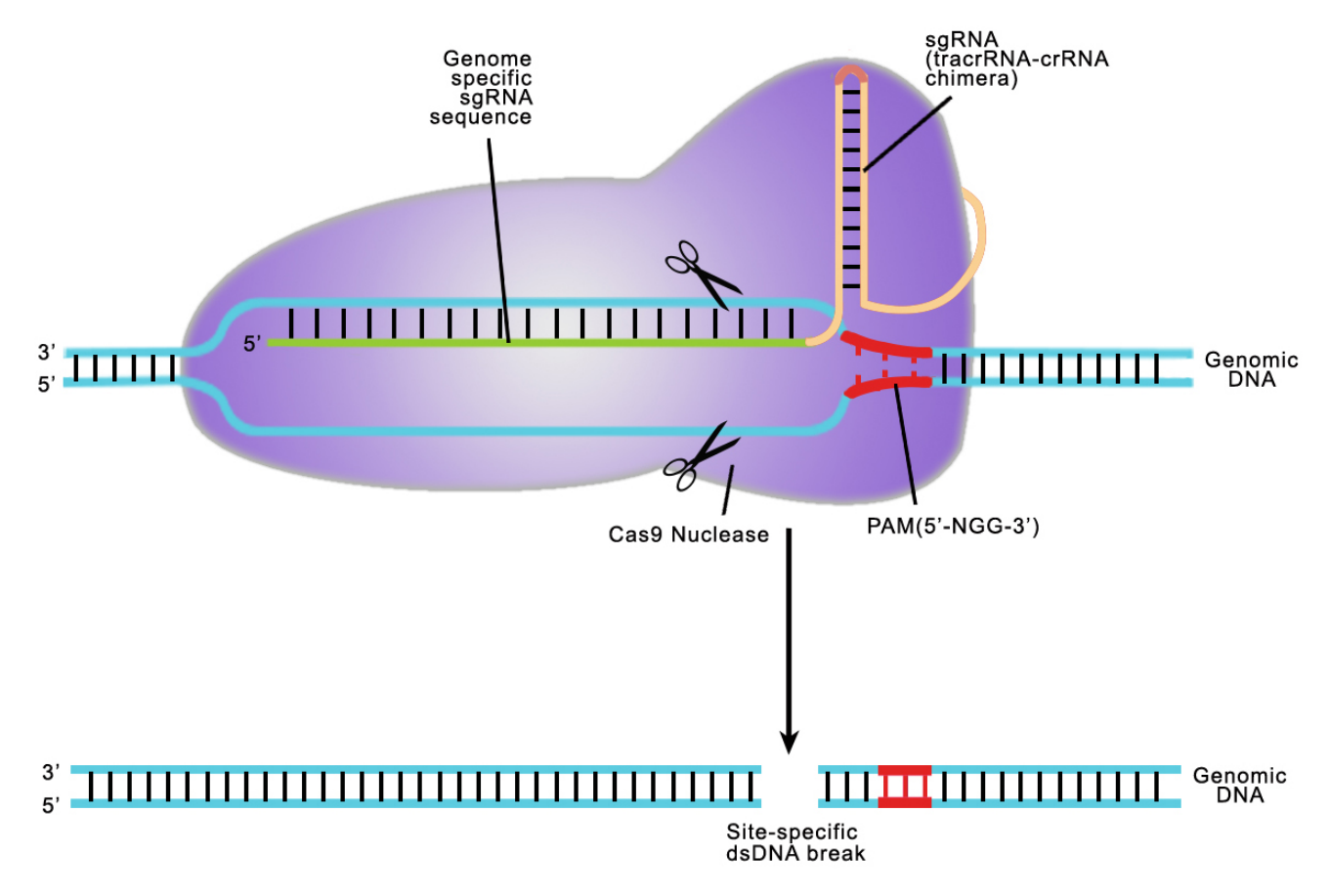
For a 3D view of the Cas9 protein, binding trRNA, crRNA and a dsDNA see:

http://www.rcsb.org/pdb/explore/explore.do?structureId=5FQ5

Hints: Switch to JSMol viewer. *First show "colour by molecule type" and "cartoon" to g*
of chromsomal DNA, tracrRNA and crRNA. Then give try other styles.

Genome Editing - Simplified Cas9-sgRNA editing system

- In nature, cas9 carries two RNAs: tracker RNA (trRNA) and crRNA which base-pairs with a sequence on trRNA.
- To make a functional CRISPR, we can synthesize chimeric RNAs (sgRNA) containing both trRNA and the crRNA targeted to the desired sequence.
- The sequence chosen must be adjacent to 5'NGG3', known as a PAM sequence, in the chromosome.

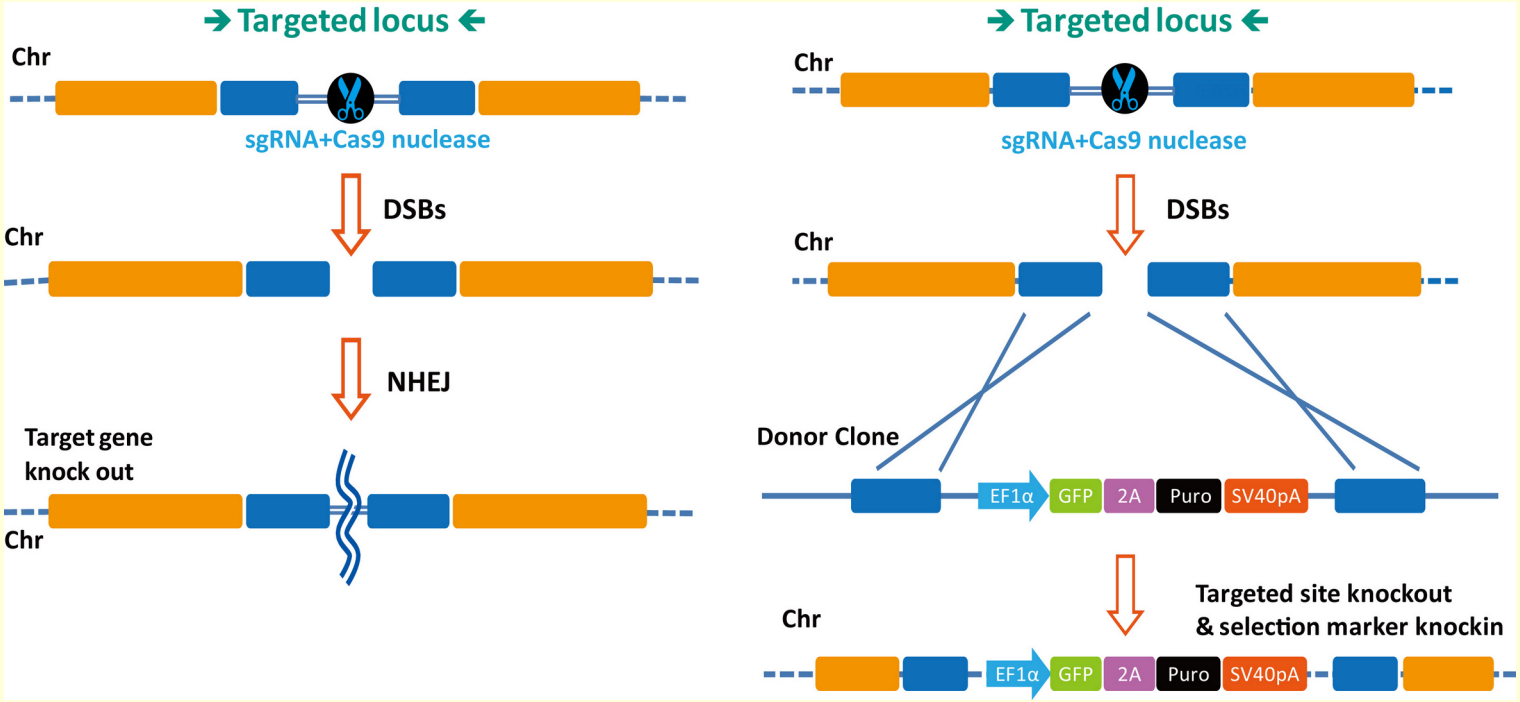


When Cas9-sgRNA are both expressed in a plant or animal cell, the CRISPR will base-pair with the complimentary chromosomal sequence and create a double-strand break.

from: http://www.genecopoeia.com/product/crispr-cas9/

Non-Homologous End-Joining (NHEJ)

Homology-Directed Repair (HDR)



Make a point mutation

DNA repair enzymes in the nucleus are error-prone. Often, several nucleotides are added or deleted when the breaks are joined.

Insert a gene construct

Co-transfection with a construct flanked by sequences homologous to the target locus can result in a double-crossover, in which the genes from the construct are inserted at the homologous site.

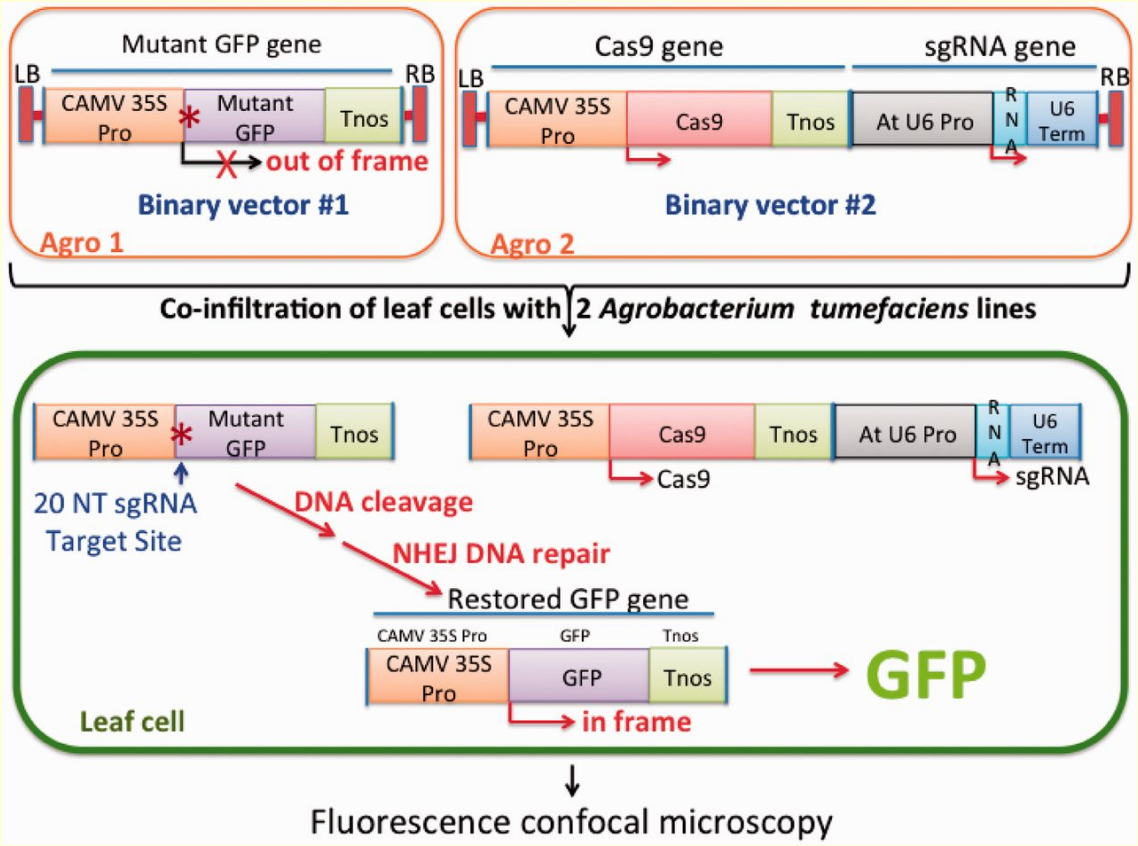
from: http://www.genecopoeia.com/product/crispr-cas9/

Genome Editing in Arabidopsis

Two constructs:
Agro1 has a non-functional GFP gene.
Agro2 has Cas9 and a sgRNA homologous to the mutant site in GFP.

Hypothesis: Co-transfection with two Agrobacterium strains containing each construct will generate mutations in the GFP gene. A mutation that restores the reading frame should make GFP functional. Only cells getting both constructs will have repaired GFP genes.

(For sgRNA gene, the Arabidopsis ubiquitin promoter and terminator were used.)



Fluorescence should only be seen in cells in which the reading frame of GFP has been restored.

Jiang et al. Nucl. Acids Res. (2013) 41 (20): e188. doi: 10.1093/nar/gkt780

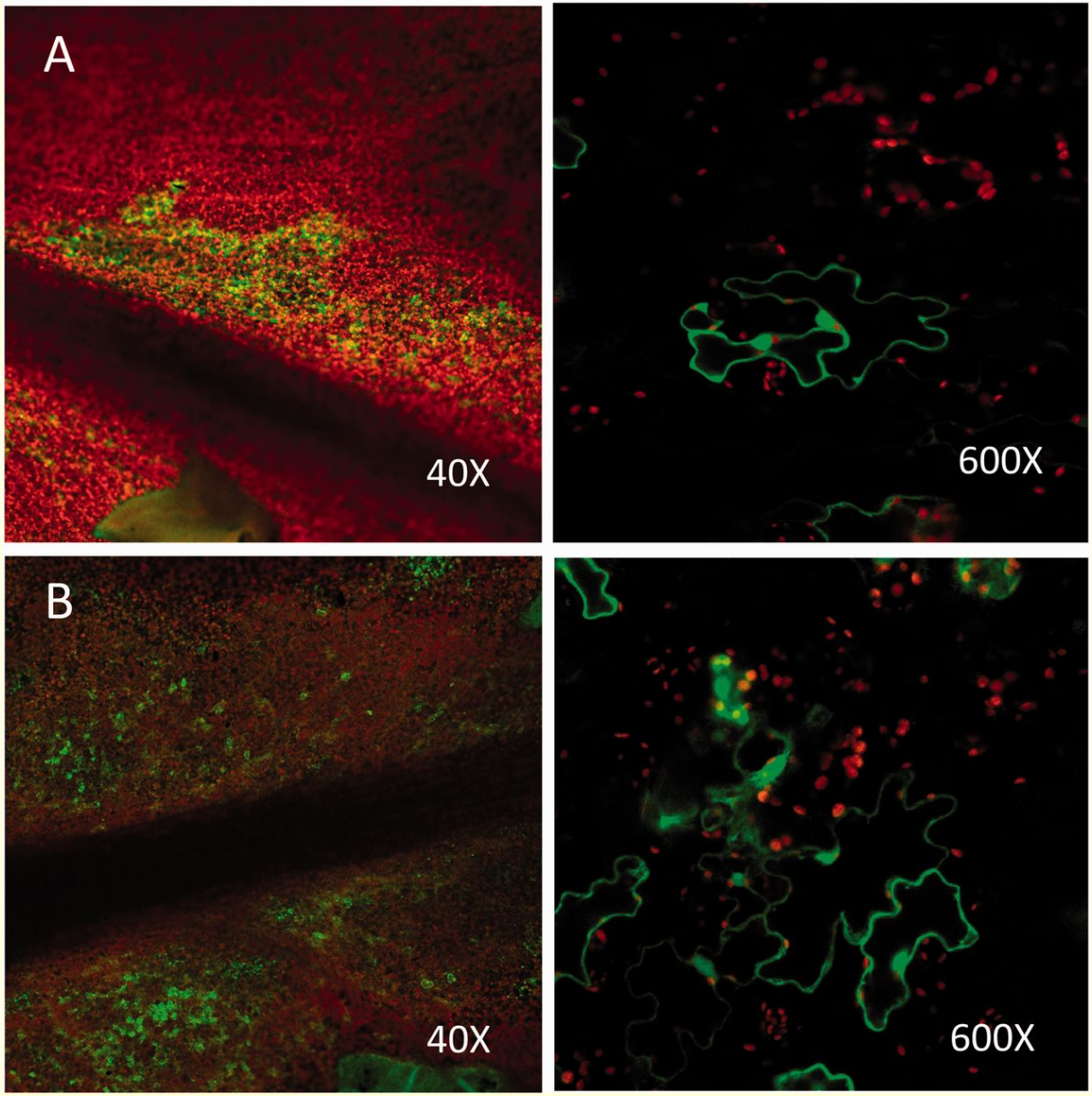
Genome Editing in Arabidopsis

A: Leaves infiltrated with wild-type GFP construct.

B: Leaves infiltrated with Agro1 and Agro2 constructs.

Red fluorescence is from chlorophyll. Green fluorescence is from GFP.

Leaves are shown 48 hr. after infiltration with Agrobacterium.



Jiang et al. Nucl. Acids Res. (2013) 41 (20): e188. doi: 10.1093/nar/gkt780

Genome Editing in Arabidopsis and Tobacco

Experiment was done in both Arabidopsis and tobacco.

1. Cut out leaf areas exhibiting fluorescence.

2. Extract DNA.

3. Cut with ApaLI (GTGCAC) to eliminate genes in which no editing has occurred.

4. Amplify presumptive mutant genes using primers 125 bp upstream and 125 bp downstream of sgRNA target site.

5. Sequence PCR fragments.

Arabidopsis: mutations in 12 out of 25 sequenced clones

	In/Del	Freq.
CATGGAGCGCTTCAAGGTGCACATGGAGGACTAGTAAAGGAGAAGAAC		
CATGGAGCGCTTCAAGGTCCCCATGGAGGACTAGTAAAGGAGAAGAAC	0 (-3,+3)	1x
CATGGAGCGCTTCAAGGTGCACAATGGAGGACTAGTAAAGGAGAAGAAC	+1	1x
CATGGAGCGCTTCAAGGTGCA...GGAGGACTAGTAAAGGAGAAGAAC	-3	4x
CATGGAGCGCTTCAAGGTGCA.....AGGACTAGTAAAGGAGAAGAAC	-5	3x
CATGGAGCGCTTCAAGGTGC......AGGACTAGTAAAGGAGAAGAAC	-6	2x
CATGGAGCGCT..............GACTAGTAAAGGAGAAGAAC	-17	1x

Tobacco: mutations in 15 out of 28 sequenced clones

	In/Del	Freq.
CATGGAGCGCTTCAAGGTGCACATGGAGGACTAGTAAAGGAGAAGAAC		
CATGGAGCGCTTCAAGGTGCACATGGAAGGACTAGTAAAGGAGAAGAAC	+1 (+1)	1x
CATGGAGCGCTTCAAGGTGCACATGCAGGACTAGTAAAGGAGAAGAAC	0 (-1,+1)	1x
CATGGAGCGCTTCAAGGTGCA...GGAGGACTAGTAAAGGAGAAGAAC	-3	4x
CATGGAGCGCTTCAAGGTG....TGGAGGACTAGTAAAGGAGAAGAAC	-4	1x
CATGGAGCGCTTCAAGGTGCA....GAGGACTAGTAAAGGAGAAGAAC	-4	1x
CATGGAGCGCTTCAAGGTGC......AGGACTAGTAAAGGAGAAGAAC	-5	3x
CATGGAGCGCTTCAAGGTGC......AGGACTAGTAAAGGAGAAGAAC	-6	3x
CATGGAGCGCTTCAAGGTGCA..............AGGAGAAGAAC	-16	1x

blue - sgRNA target site
green - inserted nucleotides
red - PAM site

Jiang et al. Nucl. Acids Res. (2013) 41 (20): e188. doi: 10.1093/nar/gkt780

Genome Editing in Poplar (Populus tomentosa)

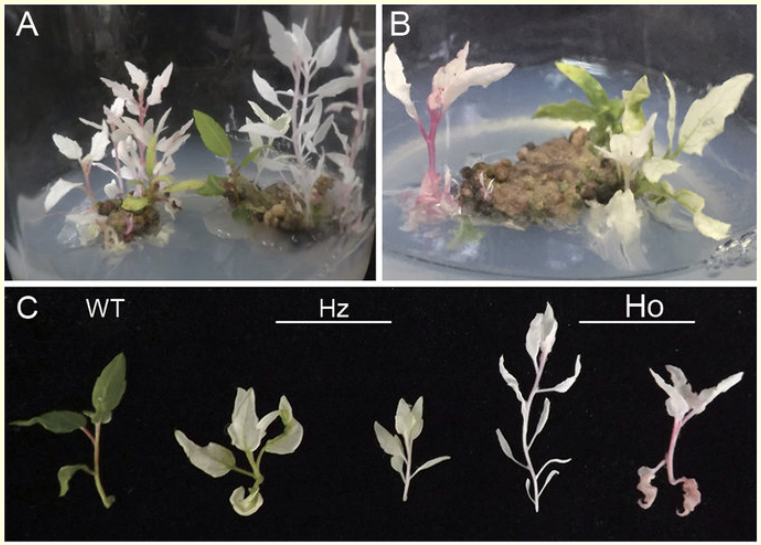
pto - phytoene desaturase. Required for chlorophyll biosynthesis.

CRISPR-Cas9 used to disrupt pto genes.

Result: Mutant albino phenotype.

A , B - regenerated plants in which pto has been knocked out.

C - WT: wild type
Hz - heterozygous mutant
Ho - “homozygous” (biallelic) mutant



Biallelic vs. homozygous
All are T0 plants ie. primary transformants. Therefore, plants that are homozygous for the albino phenotype have had both WT alleles disrupted independently. It is therefore more accurate to call these plants biallelic.

Fan D et al. (2015) Efficient CRISPR/Cas9-mediated Targeted Mutagenesis in Populus in the First Generation. Scientific Reports 5, Article number: 12217

Genome Editing - Efficiency

Target gene	No. of plants examined	No. of plants with mutations	Mutation rate (%)	Putative homozygous		Putative heterozygous	
				Number	%	Number	%
PtoPDS	59	30	51.7	28	93.2	2	6.7
CK	42	0	0	ND	ND	ND	ND

CK, Empty vector. ND, Not determined.

This table illustrates the efficiency of genome editing with CRISPR. Note that of plants with mutations, **28/30 plants had the mutation in the homozygous stat**e ie. both chromosmal copies were mutated.

These would have been independently-generated mutations ie. different mutations in each copy of the locus, both giving the mutant phenotype.

These results are in stark contrast to transformation, in which T0 transformants are always hemizygous.

Genome Editing - Off-target mutations

Rationale: Every biological system has an error rate.

It's important to determine how often CRISPR-induced mutations affect sites other than the ones to which they are targeted by the sgRNA.

Strategy:

- We can identify sites in the genome that are similar to the sgRNA, and are therefore most likely to be modified by the CRISPR, in addition to the target site.
- By comparing the mutation rate at these sites with the mutation rate at the target site, we get an estimate of the off-target mutation rate.

Genome Editing - Off-target mutations

Finding potential off-target sites:

1. Use the sgRNA to search the Soybean genome using BLASTN, E=5 (ie. allow hits such that we expect 5 hits that good, just by random chance alone).

2. Choose hits that have
- 2 - 6 mismatches with the original 23 nt sgRNA
- the PAM motif (NGG) at the 3' end of the sequence

3. Find the original sites corresponding to these hits in the Soybean genome.

4. Design PCR primers to amplify the sequences containing these off target sites.

Genome Editing - Off-target mutations

Sample BLASTN hit from soybean genome:

```
>NC_016093.2 Glycine max cultivar Williams 82 chromosome 6, Glycine_max_v2.0,
whole genome shotgun sequence
Length=51416163

 Score = 44.9 bits (23),  Expect = 1e-04
 Identities = 23/23 (100%), Gaps = 0/23 (0%)
 Strand=Plus/Plus

Query  1         GGCATGGTGCGGTCTATGAGTGG  23
                 |||||||||||||||||||||||
Sbjct  15084883  GGCATGGTGCGGTCTATGAGTGG  15084905


 Score = 29.5 bits (15),  Expect = 4.9
 Identities = 19/21 (90%), Gaps = 0/21 (0%)
 Strand=Plus/Minus

Query  3        CATGGTGCGGTCTATGAGTGG  23
                 |||| |||||||| ||||||
Sbjct  4613603  CATGTTGCGGTCTAAGAGTGG  4613583
```

perfect match with target gene, Met1-04

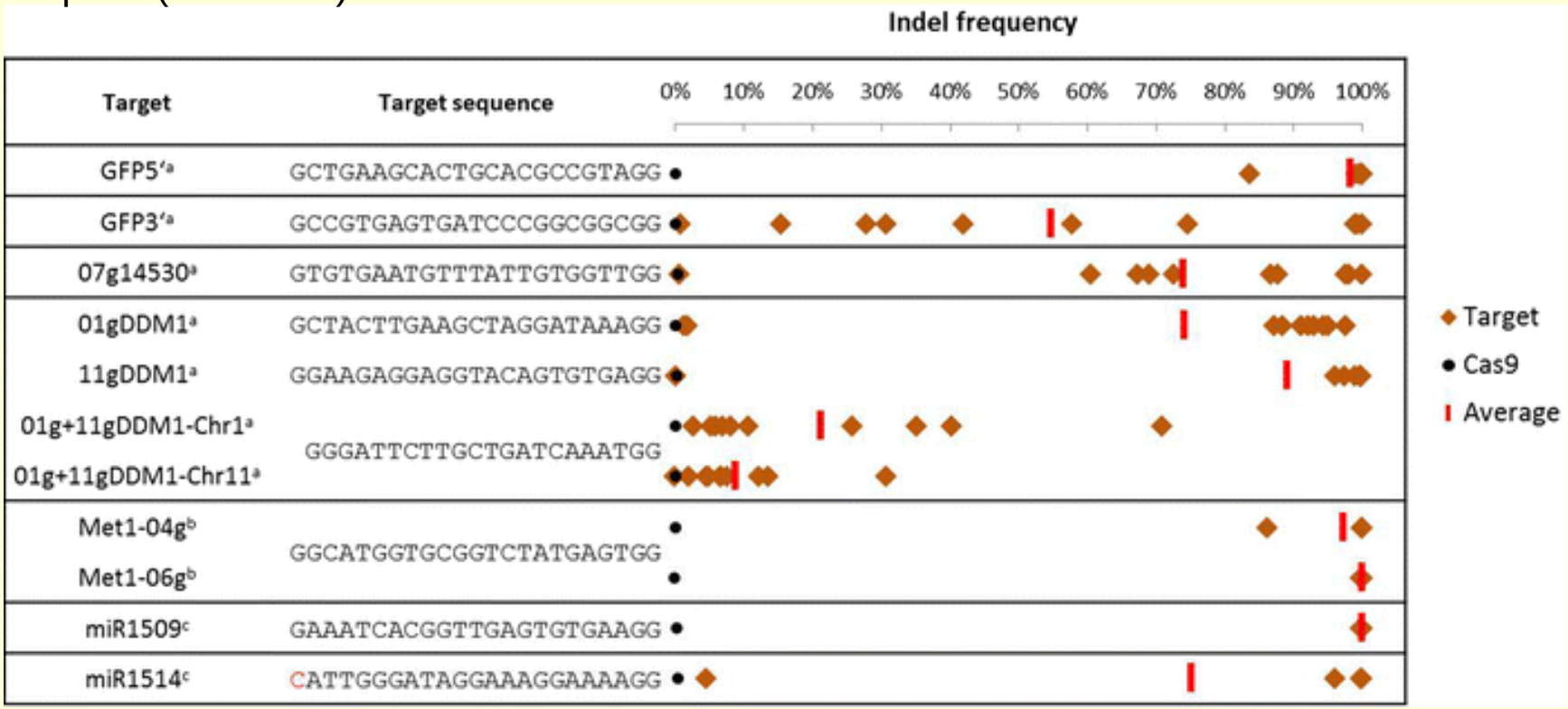
potential off-target site

potential PAM site

Genome Editing - Off-target mutations

Rate of on-target mutations (control)

1. Soybean cotyledonary explants were transformed with CRISPR constructs for 5 soybean target genes. Plants transgenic for GFP were also tested.
2. In many independently-transformed events* (explants), the sites were amplified by PCR and sequenced (~4000 - 6000 reads per event).
3. Results show the percentage of reads with insertions or deletions for a given explant (diamonds).

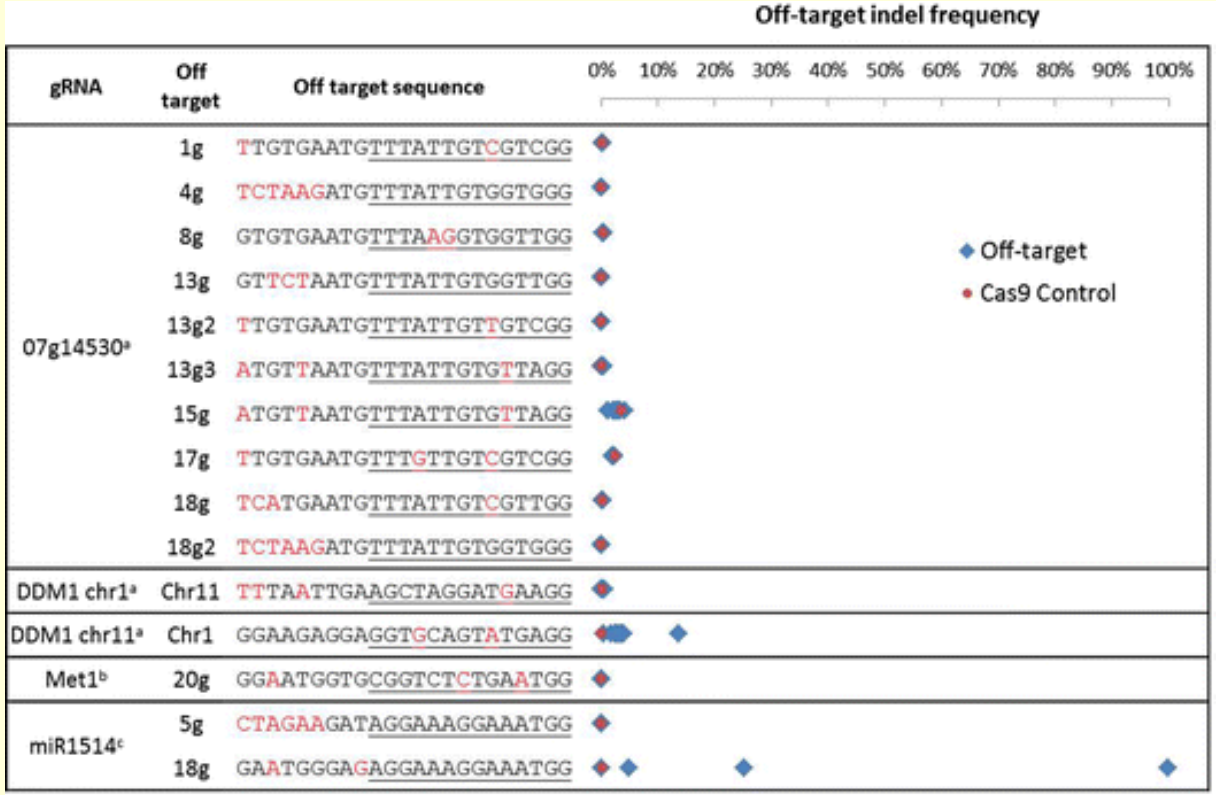


*Each explant may have many independently transformed cell lines. They are not necessarily derived from single transformed cells. Different cell lines could therefore have different mutations.

Genome Editing - Off-target mutations

Rate of off-target mutations

1. Soybean cotyledonary explants were transformed with CRISPR constructs for 4 soybean genes.
2. In many independently-transformed explants, the sites were amplified by PCR and sequenced.
3. Results show the percentage of reads with insertions or deletions for a given explant (diamonds).



Genome Editing - Off-target mutations

Conclusions

- Not all constructs gave off-target mutations
- For those constructs that did give off-target mutations, the frequency of mutations was usually very low, compared to on-target mutations.

Although off-target mutations have been observed to occur at a low rate in a large number of studies, mutant Cas9 genes have been developed that give even lower frequencies of off target mutations.

How does genome editing differ from mutatgenesis?

GENETICS	mutagenesis by chemicals or radiation	CRISPR mutations
genomic effects	large number of loci affected	single targeted locus (low level of off-target mutations)
genic effects	any possible site within the gene	narrow window targeted by sgRNA, usually ~20 bp
types of mutations	base substitutions, insertions, deletions, chromsome breaks, aneuplioidies	point mutations, small insertions/deletions
primary mutant plants	m_0 - hemizygous	m_0 - usually biallelic
dominance	usually recessive	usually recessive
identification of mutant locus	genetic crosses followed by chromosome walk	already know the identity of the gene; locus can be determined if genome is sequenced

How does genome editing differ from mutatgenesis?

PLANT BREEDING	mutagenesis by chemicals or radiation	CRISPR mutations
detection of mutation	m_1 or later requires selfing to generate homozygotes	m_0, since mutation is usually homozygous (biallelic)
plant vigor	most mutations off target, so plants are often sickly	single gene affected, so fewer unintended phenotypic effects
number of progeny screened	very large number, because mutation is random	most progeny have a mutation
back crosses	many generations needed to eliminate off-target mutations	may not need to do back crosses
screening	by pheonotype	by PCR

Conclusions:

- Genome editing reported so far in at least many plant species including Arabidopsis, tobacco, sorghum, rice, maize, Marchantia, tomato, soybean, melon and poplar
- CRISPR is not a gene delivery method. You still need to use *Agrobacterium*, biolistics, protoplasts or some other method for getting the construct into cells
- Genome editing is extremely efficient.
- No general way to tell if a plant has undergone genome editing.
  - With transgenics, you could always probe with a foreign sequence eg. 35S promoter, nptII coding sequence
  - Mutations generated by genome editing are indistinguishable from natrually-occuring mutations.
- How can/should genome editing in crops be regulated? Can it be regulated?

Genome edited crops in commercial production

- Very few gene edited crop varieties had been approved for production, although many are undergoing regulatory approval in many countries

- USA
  - Cibus - tolerant to sulfonylurea herbicides
  - Calyxt - high oleic acid soybean oil for food products

References

Doudna JA, Charpentier E (2014) The new frontier of genome engineering with CRISPR-Cas9. Science 346 DOI: 10.1126/science.1258096

Fan D et al. (2015) Efficient CRISPR/Cas9-mediated Targeted Mutagenesis in Populus in the First Generation. Scientific Reports 5, Article number: 12217 doi:10.1038/srep12217

Horvath P and Barrangou R (2010) CRISPR/Cas, the Immune System of Bacteria and Archea. Science 327:167-170 doi: 10.1126/science.1179555.

How CRISPR Works. http://www.genecopoeia.com/product/crispr-cas9/

Jiang W, Zhou H, Bi H, Fromm M, Yang B, Weeks DP (2013) Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice
Nucl. Acids Res. (2013) 41 (20): e188. doi: 10.1093/nar/gkt780

Menz J et al. (2020) Genome Edited Crops Touch the Market: A View on the Global Development and Regulatory Environment. Frontiers in Plant Science 11:586027. doi:10.3389/fpls.2020.586027.

Quétier F (2015) The CRISPR-Cas9 technology: Closer to the ultimate toolkit for targeted genome editing. Plant Science 242 p65-76. doi:10.1016/j.plantsci.2015.09.003

from: http://www.genecopoeia.com/product/crispr-cas9/