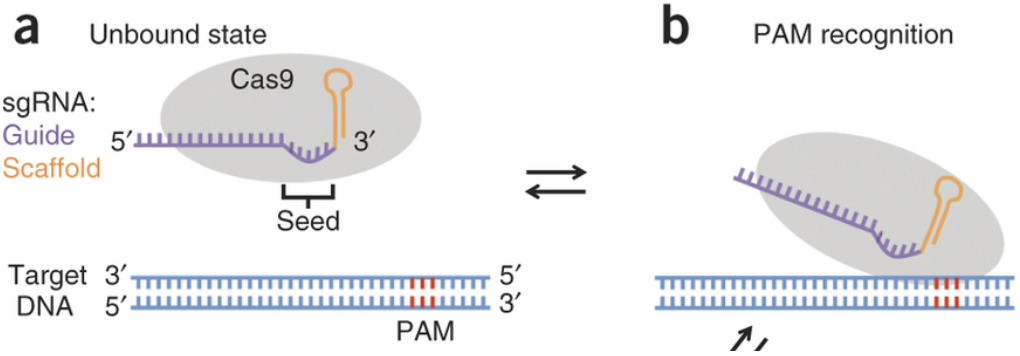
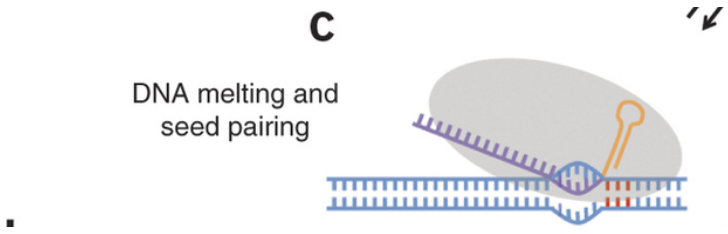
Workshop prepared by Dr Nicola Patron_ Head of Synthetic Biology at The Sainsbury Lab, Norwich.

PART 1 – SELECTING GUIDES

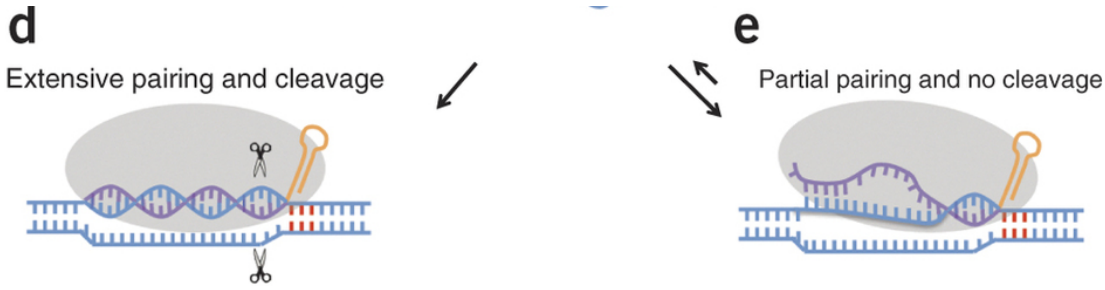
a) The single guide RNA (sgRNA) consists of a guide sequence (purple) linked to a structured scaffold sequence (orange). It forms a complex with Cas9 (grey oval). The 6-8 base pair region of the guide closest to the scaffold is known as the 'seed'.



b) The sgRNA-Cas9 complex engineered from **Streptococcus pyogenes** scans DNA until it recognizes an "NGG" Protospacer Adjacent Motif sequence (PAM, in red). **C)** After PAM recognition the complex probes the seed region to determine complementarity. If there is a match the complex will attempt to form a duplex along the entire region of the guide.



d) If there is a match, pairing will continue the Cas9 will cleave both DNA strands. Usually the cut is made 3 bases pairs of the PAM. **e)** If there is insufficient pairing the complex will be unable to cleave.



Images from Wu et al (2014) Nature Biotechnology 32, 670–676

Cas9 expression can be driven from a tissue-specific, inducible or strong constitutive promoter from a gene transcribed by RNA polymerase II (RNAPol-II).
RNAPol-II transcripts however, commonly undergo 5' and 3' end modifications and are not commonly used to drive expression of the sgRNA.

RNA polymerase III (RNAPol-III) dependent promoters are typically used to transcribe sgRNAs. Their suitability is primarily due to their precise transcriptional start site. RNAPol-III promoters such as U6 and U3 are generally used to transcribe sgRNAs

For U6 promoters (e.g. U6-26 from *Arabidopsis thaliana*) the start of transcription is generally a “G”. In order for every base in the **guide** to have a cognate base in the DNA **target,** both must begin with a “G”.

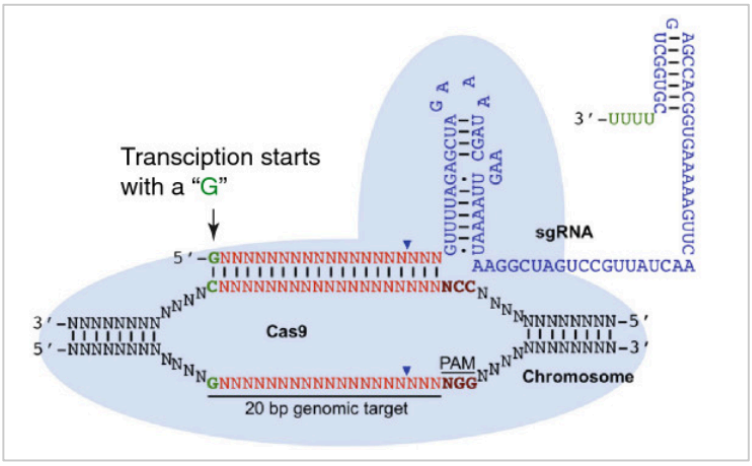
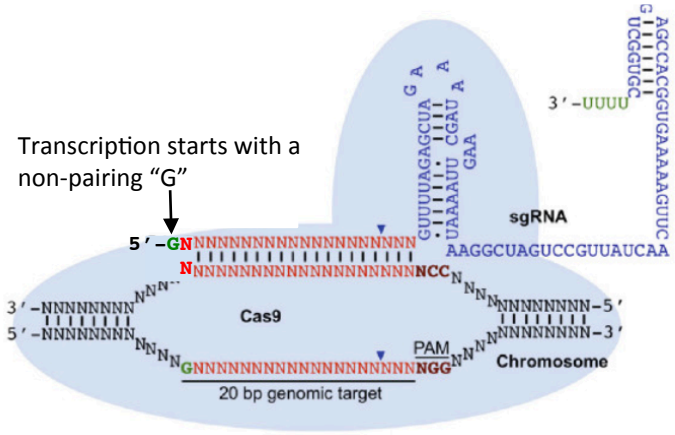


Image from Belhaj et al (2013) Plant Methods 9:39

We and others have noted that the guide does not need to start with a “G”. However, the transcriptional start site (“G”) must be preserved. Therefore, the guide must have an additional non-pairing “G” at the 5' end. We have successfully induced cleavage at targets that do not start with a “G” using an sgRNA with a additional 5’ non-pairing “G”.



In this practical you will be given two similar sequences and you will pick targets at which to induce mutations/deletions in (a) only one of these (b) both of these.

ACTIVITY 1. IDENTIFY ALL POSSIBLE TARGETS IN BOTH SEQUENCES

Open **CTest sequence1.gb** and **CTest sequence2.gb** in **Vector NTI**

NOTE: Although these sequences were provided by Dr Patron, a user can provide any sequence of interest. The search parameters will be the same.

Starting with **CTest sequence1.gb** click in the window that displays the sequence information.

In the "Edit" menu select "Find Sequence" or use the Find Sequence button 

In this Find Sequence pop-up window select "match de-generate nucleotides" this will allow you to type in any bases from the IUPAC notation:

Symbol[2]	Description	Bases represented				
A	Adenine	A				1
C	Cytosine		C			
G	Guanine			G		
T	Thymine				T	
U	Uracil				U	
W	Weak	A			T	2
S	Strong		C	G		
M	aMino	A	C			
K	Keto			G	T	
R	puRine	A		G		
Y	pYrimidine		C		T	
B	not A (**B** comes after A)		C	G	T	3
D	not C (**D** comes after C)	A		G	T	
H	not G (**H** comes after G)	A	C		T	
V	not T (**V** comes after T and U)	A	C	G		
N or **-**	any **N**ucleotide (not a gap)	A	C	G	T	4

- Search the test sequence for: "**GNNNN NNNNN NNNNN NNNNN NGG"**

- When you find a match, keep the desired sequence selected in the sequence pane and select "**New – Add Feature to F-Map**" from the "**Edit**" menu. [Or use the Add new feature button: ]

- Name the new feature "**target**" and leave the type as "**Misc. feature**". It should appear as an annotation on the graphical map.

- Repeat this search and annotation procedure, searching on the forward and reverse strands of both test sequences until all possible targets are annotated.

- Repeat again, searching for: **"NNNNN NNNNN NNNNN NNNNN NGG"**

- Evidence from the literature suggests that it may be best to avoid pyrimidines in the seed region. (**Note** - we are not yet certain if this holds true in plants – we have made more than a few pyrimidine-rich sgRNAs that have good activity but we do not have systematic data).

- Repeat your search for
NNNNN NNNNN NNNNR RRRRR NGG 3'

- Edit the name of any guides that match to "**target-PR**" (PR-for Purine-Rich).

- Save both **CTest sequence1.gb** and **CTest sequence2.gb** sequences into the local Vector NTI database.

ALTERNATIVE - USING AN ONLINE TARGET SELECTOR:

Open a web-browser and visit the following page:

https://www.dna20.com/eCommerce/cas9/input

- Select *Arabidopsis thaliana* from the species menu.

- Click the button for "wild-type Cas9"

- Click "Target a Sequence"

- Copy and paste the sequence for **Test sequence 1** into the drop-down box.

- Click "search".

Note that the scoring matrix at this site is proprietary however, part of the score is related to the number of possible off-targets in the Arabidopsis genome and part is due to the composition of the target. Different on-line tools use different scoring matrices – some of them are described in the associated publications.

These on-line tools are excellent unless you are trying to select pairs of guides that target multiple members of a family (very common for plant scientists!). They are also not particularly helpful if your species of interest is not listed. Even though you may be able to paste in a sequence of interest, a substantial part of the score will come from the number of possible off-targets in the genome.

ACTIVITY 2. SELECT SPECIFIC TARGETS

- Select "**AlignX - Open new alignment window**" from the "**Align**" menu.

- Select the "add sequence from local database" icon: 

- Drag two sequences **CTest sequence1.gb** and **CTest sequence2.gb** into the top-left pane of the AlignX program. [You may need to resize the database and Align X windows so you can see both panes].

- Select both sequences (click on each sequence with the Shift Key held down or used "Cntrl-A"

- Select "**Align Selected Sequences**" from the **Align** window.

- Wait for the alignment to appear in the bottom window.

You are now going to display the targets that we annotated previously on the multiple sequence alignment so that we can see which guides cut both sequences and which will distinguish between the sequences.

- In the top right window where your sequences are listed, click on the  box at the side of **Test sequence1.gb** then click to expand feature map. Right click on “Misc feature” and select “Show all features for ‘sequence name’ of misc. feature type”. Repeat for **Test sequence 2.gb**

- You will now see your annotated guides displayed on the alignment below.

- Identify two guides that:
  (a) Cut sequence **Test sequence1.gb** but NOT **Test sequence 2.gb**
  (b) Cut sequence **Test sequence2.gb** but NOT **Test sequence 1.gb**
  (c) Cut sequence **Test sequence1.gb** AND **Test sequence 2.gb**

PART 2 – MAKING THE CONSTRUCT

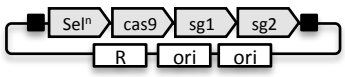
We will now design a construct that aims to delete the region of DNA between two targets in **Test sequence1.gb** *and* **Test sequence 2.gb** First, we need to understand how Type IIS mediated assembly works.

https://www.youtube.com/watch?v=xusiavAC_Xk&list=PLvzzMEb3Zsn-n-ItduNGJzghAJsgsnd4Q

These are the materials provided in the Golden Gate Plant CRISPR Kit. You will receive the annotated .gb files for these sequences by email.

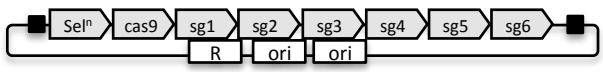
	1	2	3	4	5	6	7
A	pUAP1 Universal Level 0 Acceptor	pAGM8031 level M acceptor position 1	pICH50881 level M end-link 2	pICH51266 CaMV35s Pro + 5U GGAG-AATG	pICSL90002 AtU6-26 Pro + 5U GGAG-ATTG	pICSL90010 sgRNA template ATTG-CGCT	
B	pICH47732 level 1 acceptor position 1 forward	pAGM8043 level M acceptor position 2	pICH50892 level M end-link 3	pICSL12009 ZmUbi Pro + 5U GGAG-AATG	pICSL80037 nptII CDS AATG-GCTT	pICH41414 CaMV35s 3U + Ter GCTT-CGCT	
C	pICH47742 level 1 acceptor position 2 forward	pAGM8055 level M acceptor position 3	pICH50900 level M end-link 4	pICSL12014 OsAct Pro + 5U GGAG-AATG	pICSL80036 hptII CDS AATG-GCTT	pICH41421 AtuNos 3U + Ter GCTT-CGCT	
D	pICH47751 level 1 acceptor position 3 forward	pAGM8067 level M acceptor position 4	pICH50914 level M end-link 5	pICSL12015 AtUbi10 Pro + 5U GGAG-AATG	pICH42222 Bar CDS AATG-GCTT	pICH41432 AtuOcs 3U + Ter GCTT-CGCT	
E	pICH47761 level 1 acceptor position 4 forward	pAGM8079 level M acceptor position 5	pICH50927 level M end-link 6	pICH87633 AtuMNos Pro + 5U GGAG-AATG	pICH43844 Bar+intron CDS AATG-GCTT	pICH72400 AtuG7 3U + Ter GCTT-CGCT	
F	pICH47772 level 1 acceptor position 5 forward	pAGM8081 level M acceptor position 6	pICH50932 level M end-link 7	pICH85281 AtuMas Pro + 5U GGAG-AATG	pICSL90004 Cas9-NLS CDS AATG-GCTT	pICSL11015 LEVEL1 POS1 FAST RED	
G	pICH47781 level 1 acceptor position 6 forward	pAGM8093 level M acceptor position 7	pAGM2723-P1 level P acceptor position 1	pICH12006 CsVMV Pro + 5U GGAG-AATG	pICSL90005 Cas9-NLS CDS no stop AATG-TTCG	pICSL11017 LEVEL1 POS1 NOS-BAR-NOS	
H	pICH47791 level 1 acceptor position 7 forward	pICH50872 level M end-link 1	pICSL02208 L2 Acceptor nos-nptII-osc 35s-Cas9-nos	pICSL90003 TaU6 Pro + 5U GGAG-CTTG	pICSL50017 YFP-NLS CTAG TTCG-GCTT	pICSL11024 LEVEL1 POS1 NOS-KAN-OCS	

What will the final constructs look like?



A binary vector with a selection cassette, a Cas9 cassette and two sgRNA cassettes that target only sequence 1 R = bacterial resistance ori = origin of replication for *E. coli* / *A. tumefaciens*. Black boxes= LB/RB

Or, a larger construct with multiple sgRNAs for each target sequence to increase the chances of success.



A binary vector with a selection cassette, a Cas9 cassette and six sgRNA cassettes: 2 that target both sequence 1 and 2; 2 that target sequence 1 only; 2 that target sequence 2 only. R= bacterial resistance ori=origin of replications for *E.coli*/ and *A.tumefaciens*. Black boxes= LB/RB

1) **Choose your selection cassette.**
The kit contains two selection cassettes in **level 1 position 1:**
pICSL11017 NOS:KAN:NOS
pICSL11024 NOS:BAR:OCS
If these are not suitable for your transformation protocol, you can use the **NPTII (pICSL80037) HPTII (pICSL80036)** or **BAR (pICH42222)** level 0 standard parts, (or a part from your own collection) along with promoters and terminators parts of your own choice to assemble a suitable selection cassette in the **level 1 position 1 acceptor (piCH47732).** Several different promoters (green) and terminators (red) are provided in the kit. Alternatively you can make your own level 0 promoter and terminator parts from your species of interest. You can find a one-step assembly protocol to make Level 1 transcriptional units from level 0 parts at **http://synbio.tsl.ac.uk/golden-gate-assembly-protocol/**

Level 1 Position 1

2) **Choose your Cas9 cassette.**

The kit contains two level 0 Cas9 modules.

pICSL90004 can be assembled between any promoter (green) and terminator (red) into the level 1 position 2 acceptor (pICH47742):

pICSL90005 has no stop codon and can be assembled with a C-terminal YFP tag (pICSL50017) to check for expression and nuclear localization. When using a new Cas9 cassette or a new species it is good to check that you can see YFP in the nucleus. The YFP-tagged Cas9 retains nuclease activity.

3) **Choose your final backbone.**

Level 2 acceptors can hold up to **6** transcriptional units.

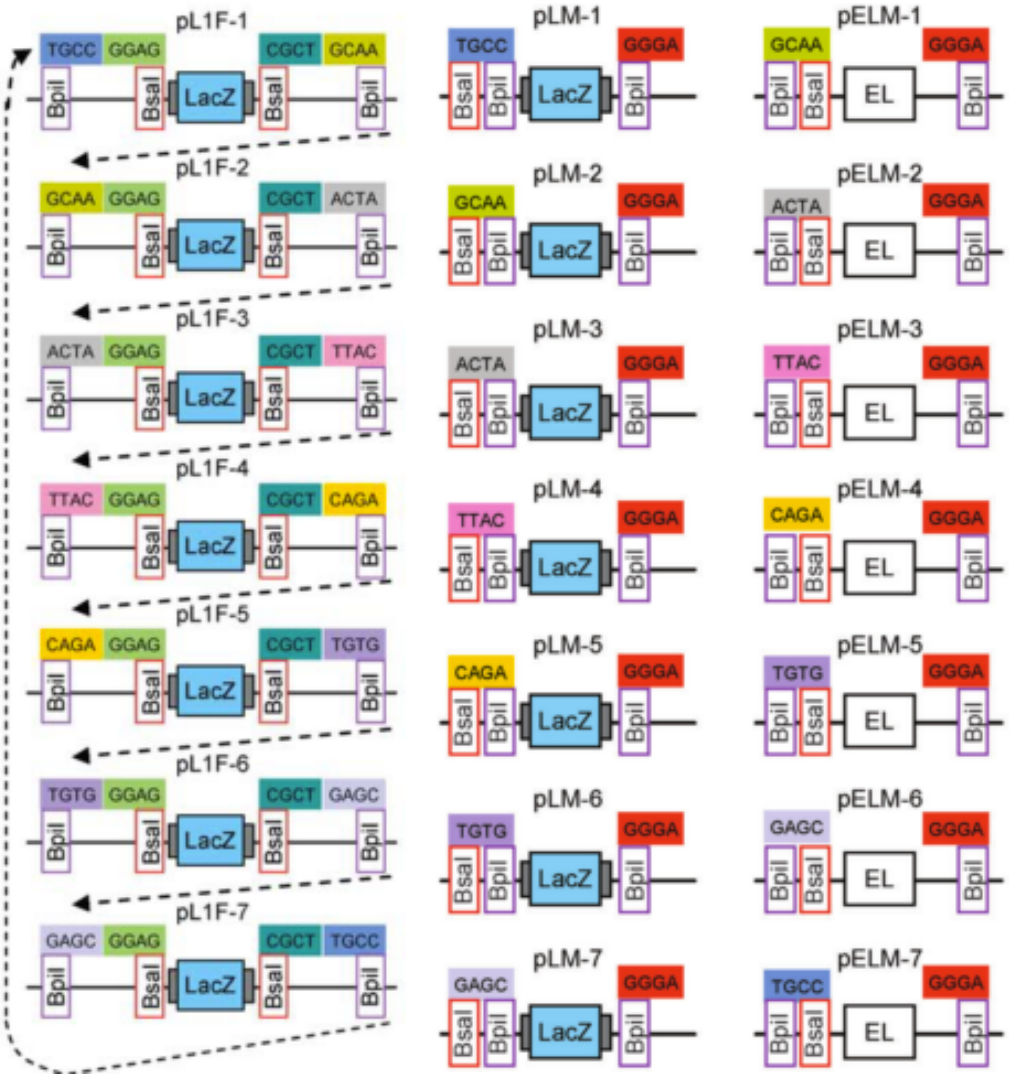
Option 1 (two sgRNAs) in a **level M position 1** (or a level 2 – not shown) vector.

Option 2 will need to be assembled in a Level **P** vector. It will have a total of 8 transcriptional units (6 sgRNAs, 1 selection cassette, 1 Cas9 cassette).

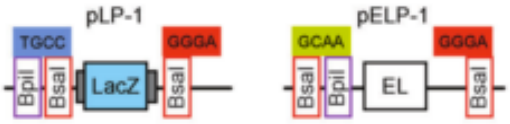
Level 1 Acceptors
Positions 1 -7 (CarbR)

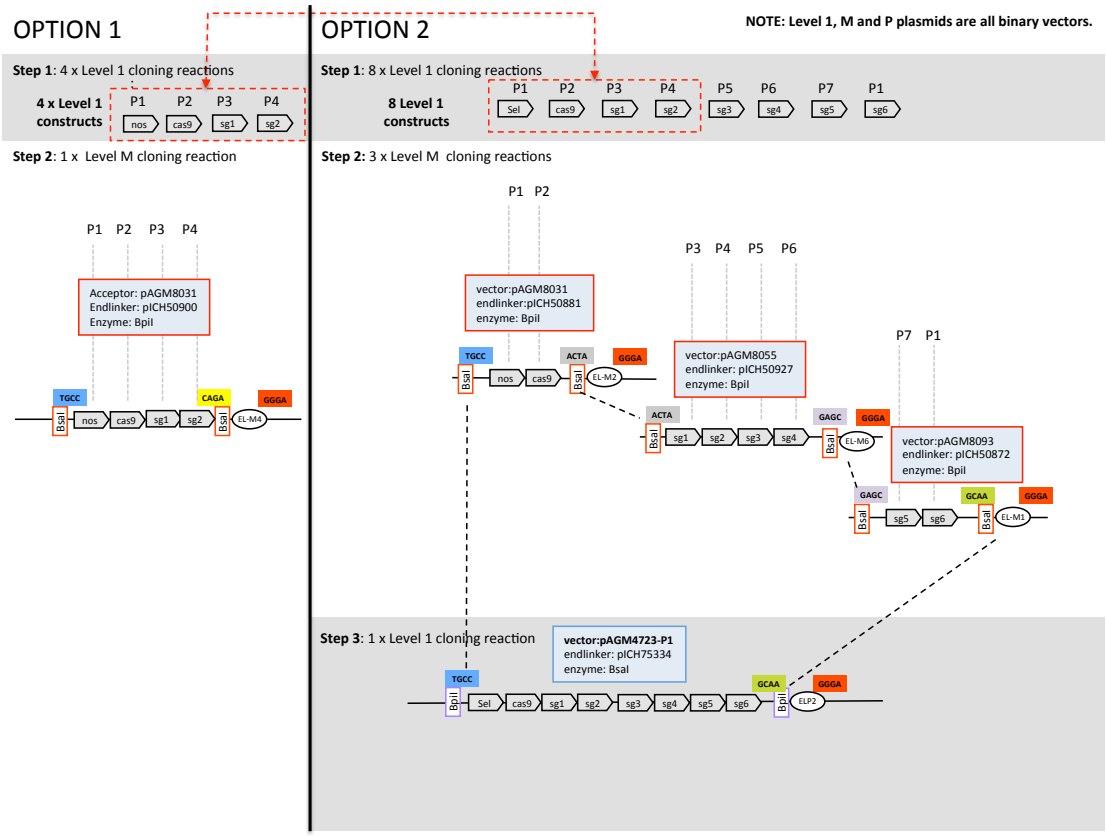
Level M Acceptors
Positions 1 -7 (SpecR)

Level M End Linkers
Positions 1 -7 (AmpR)



Level P1 acceptor (KanR) and end Linker (AmpR)




OPTION 1
OPTION 2
NOTE: Level 1, M and P plasmids are all binary vectors.
Step 1: 4 x Level 1 cloning reactions
4 x Level 1 constructs
P1 P2 P3 P4
nos cas9 sg1 sg2
Step 2: 1 x Level M cloning reaction
Acceptor: pAGM8031
Endlinker: pICH50900
Enzyme: BpiI
TGCC
CAGA
GGGA
BsaI
EL-M4
Step 1: 8 x Level 1 cloning reactions
8 Level 1 constructs
P1 P2 P3 P4 P5 P6 P7 P1
SeI cas9 sg1 sg2 sg3 sg4 sg5 sg6
Step 2: 3 x Level M cloning reactions
vector:pAGM8031
endlinker:pICH50881
enzyme: BpiI
ACTA
EL-M2
vector:pAGM8055
endlinker: pICH50927
enzyme: BpiI
GAGC
EL-M6
vector:pAGM8093
endlinker: pICH50872
enzyme: BpiI
GCAA
EL-M3
Step 3: 1 x Level 1 cloning reaction
vector:pAGM4723-P1
endlinker: pICH75334
enzyme: BsaI
BpiI
ELP2

4) Making the Level 1 sgRNA cassettes:

The sgRNA is amplified from a plasmid containing the sgRNA scaffold introducing the guide as a 5' tail in the primer. This amplicon is then assembled with a promoter in a Level 1 acceptor.

The forward primer will contain the 20 bp guide sequence specific to your chosen target as a 5′ tail.

Note that the PAM is NOT included in the sgRNA (see the figures on Page 2).

If the Arabidopsis U6-26 (**pICSL90002**) promoter is being used then the forward primer will look like this:

tgtggtctca ATTG NNNN NNNNN NNNNN NNNNN gtttaagagctatgctggaaacag
(BsaI site is in blue, **the 20 bp guide sequence** is in red, the 3′ end in black lower-case anneals to the PCR template containing the sgRNA scaffold)

If the guide does not begin with a **G** then an additional nucleotide should be included in the forward primer. The **G** must be included as this is (a) the start of transcription and (b) part of the 4 bp GoldenGate overhang by which the promoter and the sgRNA will join. The primer (and resulting guide) will be one base pair longer:

tgtggtctca ATTG **N** NNNN NNNNN NNNNN NNNNN gtttaagagctatgctggaaacag
(BsaI site is in blue, the 20 bp guide sequence is in red, the 3′ end in black lower-case anneals to the sgRNA target)

If the wheat U6 promoter (**pICSL90003**) is being used then the forward primer will look like this:

tgtggtctca CTTG (**N**) NNNN NNNNN NNNNN NNNNN gtttaagagctatgctggaaacag
(BsaI site is in blue, the 20 bp guide sequence is in red, the 3′ end in black lower-case anneals to the sgRNA target)

The reverse primer will always be the same:

tgtggtctct AGCG aaaaaaagcaccgactcggtgccac
(BsaI site is in blue, the 3′ end in black lower-case anneals to the sgRNA target)

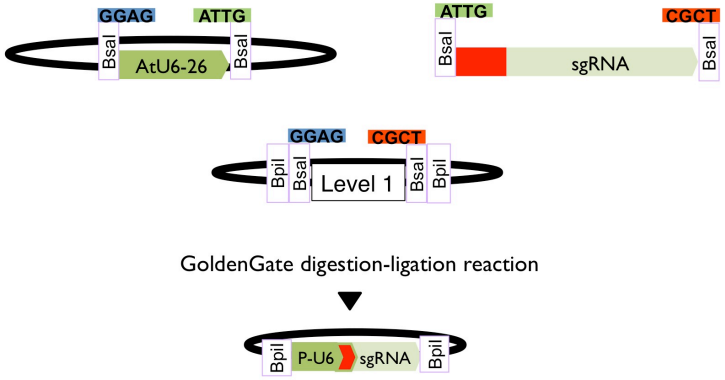
Use the forward and reverse primer pair to amplify a unique sgRNA from a suitable template with proofreading polymerase. The vector **pICSL90010** contains the sgRNA scaffold sequence. Amplification from this template with the primers pair from the previous step will result in an amplicon like this:

tgtggtctca**ATT**G(N)NNNNNNNNNNNNNNNNNNNgtttaagagctatgctggaaacagcatagcaagtttaaat
aaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgctttttt**CGCT**agagaccaca

(BsaI sites are in blue, the 20 bp guide sequence is in red)

Clean up the amplicon and quantify the DNA.

Set up a Golden Gate reaction with this PCR product as described at **http://synbio.tsl.ac.uk/golden-gate-assembly-protocol/** using the correct level 1 acceptor plasmid and a plasmid containing a small RNA promoter (e.g. pICSL0002 or 3 or a U6 promoter part specific to your species of interest). The reaction will process as shown in the figure below:



Check your clone by restriction-analysis and/or sequencing.

You can now use all of your assembled sgRNA transcriptional units in Level M reactions. You can find a one-step assembly protocol to make Level M multigene constructs from level 1 parts at **http://synbio.tsl.ac.uk/golden-gate-assembly-protocol/.** A Level P construct can then be assembled from your level M assemblies.

ACTIVITY 3:

Design a forward primer for each of the 2 guides your selected in step one for use with the AtU6-26 promoter:

_ _ _ _ _ _ _ _ _ _ _ _ _ (_) _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ g t t t a a g a g c t a t g c t g g a a a c a g

_ _ _ _ _ _ _ _ _ _ _ _ _ (_) _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ g t t t a a g a g c t a t g c t g g a a a c a g

PART 3 – SCREENING TRANSGENIC PLANTS

1. The simplest screen.

Extract gDNA from your transgenic plants and amplify the target region. As a control check that you get a clean, bright single product from WT-DNA that gives single peaks when sequenced.

ACTIVITY 4 – SELECTING PRIMERS

- Return to Vector NTI and select a pair of primers flanking your putative deleted region for each sequence. You need to select a pair of primers at least 150 bp outside of the largest possible deleted region. Your primers for each should be specific so check your alignment and choose primers from the regions in which there is the least similarity! *You can use the primer design feature of VNTI or any other progamme you feel comfortable with or design by eye; teaching good primer design is not a feature of this course!*

- Annotate the sequence with your primers. Keep the desired sequence selected in the sequence pane and select "**New – Add Feature to F-Map**" from the "**Edit**" menu. [Or use the Add new feature button: ]. Name the new feature "**primer**" It should appear as an annotation on the graphical map.

- Look for a smaller product, indicative of a deletion, in plants that contain the sgRNA-Cas9 T-DNA.

- If you only see a smaller band of the expected size you may have a homozygously edited plant. You might be able to celebrate after the next step. If you see both the WT band and the smaller band then you probably have another generation to go before you can celebrate.

- Gel-purify and direct sequence the smaller band. If you see only an edited product you can definitely celebrate. if the sequence becomes very messy at the proposed cut-site then your construct is working. Your plant is either heterozygous (one allele edited) or chimeric (Cas9 is being expressed somatically – the repair will be different in every cell). Clone the PCR product and sequence several clones (at least 10) to confirm that you have editing. If this smaller band is quite strong you probably have a good chance of obtaining a non-chimeric plant (homozygous or heterozygous) in the next generation.

- In either case your plant still has the T-DNA expressing Cas9. You probably want to segregate or cross this out so that your mutation/deletion can be shown to be heritable independent of the T-DNA. This makes it absolutely certain that you have a stable, heritable mutation and that the mutations are not the product of expression from the T-DNA insertion.

- Although you might see a deletion in the first generation this might not be the most common mutation. If you cannot find a deletion in the second generation it might be worth screening for smaller in-dels at each sgRNA. You can screen for in-dels and SNPs by sequencing or by using a *Cel*I-based Surveyor™ assay or using the digestion resistance assay described in step 3, below.

2. What is the smaller band is really, really light – too light to sequence.

ACTIVITY 4 – SELECTING RESTRICTION SITES FOR ENRICHMENT

Take note of the positions at which your amplicon will start and end (your primer positions).

Select "Restriction Analysis" from the "Analysis" menu – or click the following icon: 
- Click "Add"
- Click "Select all"
- Click "OK"
- Fill in the box to not display enzymes that cut more that 4 times.
- Fill in the start and end position of your amplicon (the positions of your primers) in the boxes under "ignore RENS cutting outside region"
- Click "OK"

- Take note of one or two restriction enzymes (preferably those that are not sensitive to methylation) that cut between your chosen sgRNAs but **not** between your primers and the sgRNAs.

- Digest an aliquot of your genomic DNA with these enzymes. Digest to completion (>3 hours in a 10 fold digestion)
- Repeat your PCR on this pre-digested and on non-digested DNA. NOTE - be careful to clean up your digestion reaction and only to add a small amount into the PCR or it may inhibit the amplification.

Your smaller band should now be much brighter and the larger WT band should be much less prominent or (hopefully) completely absent.

3. I don't see a deletion!

It is possible that only one sgRNA is working in which case you will only obtain small in-dels. You can screen for in-dels and SNPS by sequencing or by using a *Cel*I-based Surveyor™ assay or using the digestion resistance assay described in step 3, below.

ACTIVITY 5 – SELECTING MORE RESTRICTION SITES FOR ENRICHMENT

Select one or more restriction enzymes for which the recognition site overlaps with the proposed cut site of your sgRNA (3 bases pairs 5' of the PAM). Try to pick an enzyme that only cuts once within the amplicon (you can change the position of the primers that you use to amplify the cut-site if necessary). It will not be possible to find an enzyme for every guide.

Amplify from non-digested gDNA. Digest the PCR product with the chosen enzyme. The WT product should digest as expected. Any product product that does not digest might have had the recognition site disrupted. This un-cut product can be sequenced to verify this.

You can probably expect 5-20% of your first generation to show some activity (deletion or mutation). Success will depend on your sgRNAs. To increase the chances of success try making and delivering more than one construct with two or more sgRNAs to each target in each. The rate of heritability, in our experience, is species-dependent.