Instructions for characterizing insertion sites by PCR

By Nina Ivanova and Ru Zhang
Updated by Martin Jonikas 11/10/15
Updated by Xiaobo Li 2/20/16

Background on mutant generation and insertion site mapping

The generation and mapping of mutants labeled LMJ.SG0182.**** are explained in detail in the article describing this collection (Li *et al.*, *Plant Cell* 2016). The generation and mapping of mutants labeled LMJ.RY0402.****** are similar in general, with some technical refinements (Li *et al.*, unpublished).** Briefly, mutants were generated by electroporation of a DNA cassette (pMJ013b for LMJ.SG0182 mutants or CIB1 for LMJ.RY0402 mutants) conferring resistance to paromomycin into *Chlamydomonas* strain CMJ030 (Zhang *et al., Plant Cell* 2014; deposited in the Chlamydomonas Resource Center as CC-4533). Strain CMJ030 was isolated from the progeny of a cross between strains $D66^+$ (Schnell and Lefebvre, *Genetics* 1993) and $4A^-$ (Tran et al., 2012). CMJ030 can grow photoautotrophically, mixotrophically, and heterotrophically; is mating type minus; has normal swimming and lipid accumulation; has high transformation efficiency; and recovers efficiently from cryogenic storage in liquid nitrogen. Insertion sites were mapped using a method called ChlaMmeSeq (*Chlamydomonas* MmeI-based insertion site Sequencing; Zhang *et al., Plant Cell* 2014), which yields 20bp of sequence flanking one or both sides of the cassette.

Please note that all the flanking sequences reported in our datasets are read from the cassette outwards. If the cassette is inserted in the same direction as the chromosome (+), the 5' side flanking sequence is the reverse complement of the genome sequence, and the 3' side flanking sequence is the same as the genome sequence. If the cassette is inserted in the opposite direction relative to the chromosome (-), the 3' side flanking sequence is the reverse complement of the genome sequence, and the 5' side flanking sequence is the same as the genome sequence.

It is critical to streak each mutant to single colonies and characterize the insertion site by PCR prior to any further work. The culture you receive may be a mixture of several strains. Additionally, due to the messiness of insertion sites, the insertion sites reported by our mapping tools are only accurate in ~75% of the mutants.

The insertion site in your mutant may be complex. As you work to characterize the insertion site in your mutant, it will be helpful to be aware of what arrangements have been seen in other mutants. Please see Figure 2 of Zhang *et al., Plant Cell* 2014 (open access) for examples of different types of insertions, and Figure 6 of the same paper for our favorite model of the mechanism of insertion, which offers the simplest explanation for all the insertion sites we have observed. Most of the insertions we have observed in our collections have one or more of the following challenging characteristics:

- The cassette may be truncated at one or both ends.
- Multiple cassettes or cassette fragments may be concatenated at the insertion site
- One or more short (typically <500bp) fragments of genomic DNA may have been inserted between the cassette and the flanking genome at the site of insertion.
- There may be a deletion or duplication of the genomic sequence at the insertion site. In the mutants we have analyzed, most deletions and duplications were small (<100bp).
- There may be a chromosomal rearrangement, such as an inversion of a segment of DNA near the insertion site.

Protocol overview

The protocol below is carried out in several consecutive steps to verify the presence of a transforming cassette in the strain of interest by PCR. Prior to carrying out the PCR, the strain should be streaked out to single colonies. First, the DNA is extracted from a colony, then it is used to do check PCR with wild-type control and locus-specific primers, and finally, cassette-genome junctions are amplified.

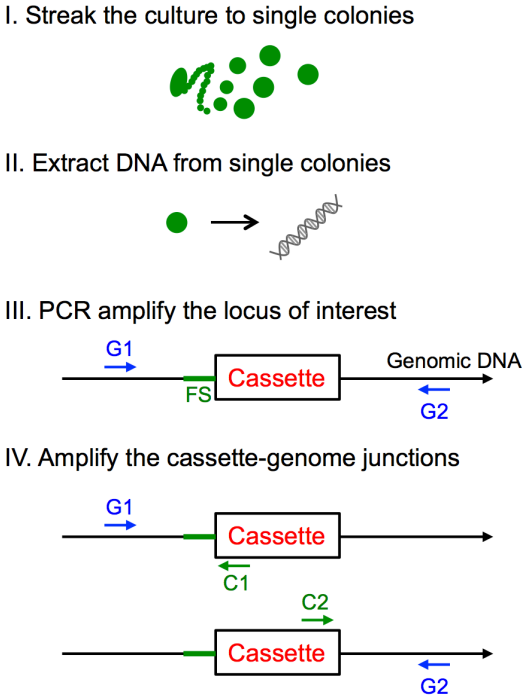

Figure 1. The four steps of the protocol are illustrated.

I. Streak the culture to single colonies

Due to the way our library is propagated, there is a chance that the culture you received is a mixture of several strains. By streaking to singles, you will ensure that you are working with a single strain.

Streak colonies to singles from the mutant and wild-type parental strain CMJ030 = CC-4533. The goal is to produce a plate where several individual cells gave rise to physically separate colonies. One technique for achieving this is as follows:

1) Using the flat end of a sterile flat wooden toothpick, transfer a clump of cells from the shipping vial to a fresh plate of TAP agar. Make a little patch out of the clump on the plate by swirling the toothpick in a circle on the surface of the agar(try not to gouge the agar). Discard the toothpick.
2) Take another toothpick, and slide it across the patch (the goal is to pick up a small number of cells), then without lifting it off the agar, rub it back and forth on the agar, such that cells get spread out on the agar. This is called a "streak". Discard the toothpick.
3) Take another toothpick, and rub it across the streak you just generated, avoiding the clump of cells you deposited in step 1- the goal is to pick up a smaller number of cells. Without lifting the toothpick off the agar, rub it back and forth over an untouched part of the agar plate. This is the region where individual colonies should grow up.

4) Let the plate grow until you see colonies. If you see round colonies that are physically separate, you were successful. If not, get help from someone who has done it before. You will need to try again, perhaps repeating step 3 several times.

II. Extract DNA from single colonies

We typically recommend analyzing multiple colonies per mutant, in case the culture you received is inadvertently a mixture several strains. You will also need at least one colony from the wild-type parental strain CMJ030 = CC-4533.

5) Grow colonies until they are at least 3mm diameter to yield enough DNA for subsequent PCR. Smaller colonies won't work well.
6) Add 50 μL of 10 mM EDTA, pH 8.0 to as many PCR tubes as DNA samples being extracted (e.g. 8 tubes for 8 samples).
7) Using a sterile pipette tip or toothpick, pick the bulk of a single colony into each tube and swirl to resuspend (leave a small part of the colony on the plate). You will need to keep track of which colonies you extract DNA from, so that you can start future experiments from what is left of the specific colony that gave the desired PCR results. After swirling, the solution in the tube should be light green.
8) Seal the tubes tightly and vortex 10 seconds on a table-top vortexer at maximum speed.
9) Boil the solution at 100°C for 10 minutes, cool at 4°C for 1 minute (this can be done in a thermocycler).
10) Vortex samples 10 seconds on maximum speed.
11) Centrifuge samples at 1000g for 1 minute.
12) The supernatant should now contain DNA. Draw the supernatant off the pellet, avoiding pellet debris, and transfer the supernatant to a fresh tube. The supernatant can be frozen at -20°C.

Example: For the PCR reactions in the example below, we extracted DNA from two wild-type colonies: WT-A and WT-B, as well as three colonies of the control mutant cMJ078 generated with the pMJ013b cassette: cMJ078-A, cMJ078-B, and cMJ078-C.

III. PCR amplify the locus of interest in wild-type and mutants

1) Design and order primers at least 1 kb away from the flanking sequence on either side of the insert. The webpage on https://chlamylibrary.org corresponding to your insertion of interest contains recommended primers. If these primers don't work, you can design others using Primer3 for primer design: http://bioinfo.ut.ee/primer3-0.4.0/. We typically design primers with target melting temperatures of 60degC. These primers will be used to amplify the region in wild-type and mutant, with the left primer being in the same direction as the genome, and the right primer being the reverse complement.

Example: for our PCR of the insertion in strain CMJ078 with an insertion on chromosome 5, we designed the following primers:

Forward:	oMJ274	GAGCAACGACCACAACACAC
Reverse:	oMJ276	CTCCTGCTGCTTTTGCTTCT

2) You will also need primers oMJ282 and oMJ284 to amplify a control locus, and later you will need primers for PCR out of the cassette. Here are the primer sequences:

Control locus:

Forward:	oMJ282	ATGCTTCTCTGCATCCGTCT
Reverse:	oMJ284	ATGTTTTACGTCCAGTCCGC

Primers binding the pMJ013b cassette and facing outwards towards the genome:

5′ end:	oMJ005	GCTGGCACGAGTACGGGTTG
3′ end:	oMJ155	GCTCGTGGAGCTCTGAATCT

Primers binding the CIB1 cassette and facing outwards towards the genome:

5′ end:	oMJ913	GCACCAATCATGTCAAGCCT
3′ end:	oMJ944	GACGTTACAGCACACCCTTG

3) Amplify the control locus and the locus of interest, using the Qiagen Taq PCR Core Kit and 1μL of the DNA extracted in Part I per 25 μL of PCR reaction. You will need to run a separate reaction for each of the DNA samples with two sets of primers: the control locus (to check DNA quality, should give strong amplification in all samples) and the cMJ078 disrupted locus. We recommend making master mixes of the PCR reagents common to all reactions, and only vary the primers and template DNA. We also recommend including a sample where the template DNA is omitted, to control for contaminated reagents. Each reaction should contain the following:

Component	**Volume for each reaction (μL)**
5x Q solution	5
10x Buffer	2.5
100% DMSO	1.25
10 mM dNTPs	0.5
Pure water	12.15
Taq	0.1
Left primer, 10 μM	1.25
Right primer, 10 μM	1.25
DNA from colony extraction	1
Total volume	**25**

4) Use the PCR conditions listed below to amplify the regions.

	Temperature °C	**Time**
once:	95	5 min
40 cycles:	95	30 s
	58	45 s
	72	2 min
once:	72	10 min
once:	10	∞

5) Run the samples on an agarose gel to visualize the amplicons. Amplification of the control locus should give a band of 1315bp in all samples containing Chlamy DNA. Amplification of the locus of interest should give a band corresponding to the distance between the primers

with wild-type template DNA, but no product or a larger product with template DNA from the mutant of interest.

6) Sequence all PCR products (excluding those across the control locus) with one or both primers to verify that the correct locus has been amplified.
    i. If no product is observed from the amplification of the control locus, there is likely a problem with your DNA extraction or PCR.
    ii. If no product is observed in the PCR across the locus of interest with wild-type template DNA, but the PCR of the control locus gave bands, one or both of the primers targeting the locus of interest are not working and need to be re-designed.
    iii. If both the wild-type and mutant yield a product of similar size for the locus of interest, sequence all bands: the mutant could have a very short insertion of a fragment of the cassette, or the locus-specific primers could be amplifying the wrong locus.
    iv. If both CMJ030 and the mutant have the same expected PCR bands (confirmed by sequencing), that mutant colony carries a wild-type version of the locus. Either the insertion site was not mapped correctly (happens in 25% of mutants), or a strain mix-up occurred, or the culture you received is a mixture of several mutants. If the desired disruption is critical for your work, you may want to repeat the procedure with additional colonies in the hopes that you are dealing with a mixture of several strains, and another colony will carry the disruption.
    v. If only CMJ030 has the expected PCR band but the mutant does not have it, or the mutant has a longer band, it is likely that the insertion site reported by the flanking sequence is correct. Continue with the next step.

Example: for our PCR of the insertion in strain CMJ078, we obtained the results shown in Figure 2.

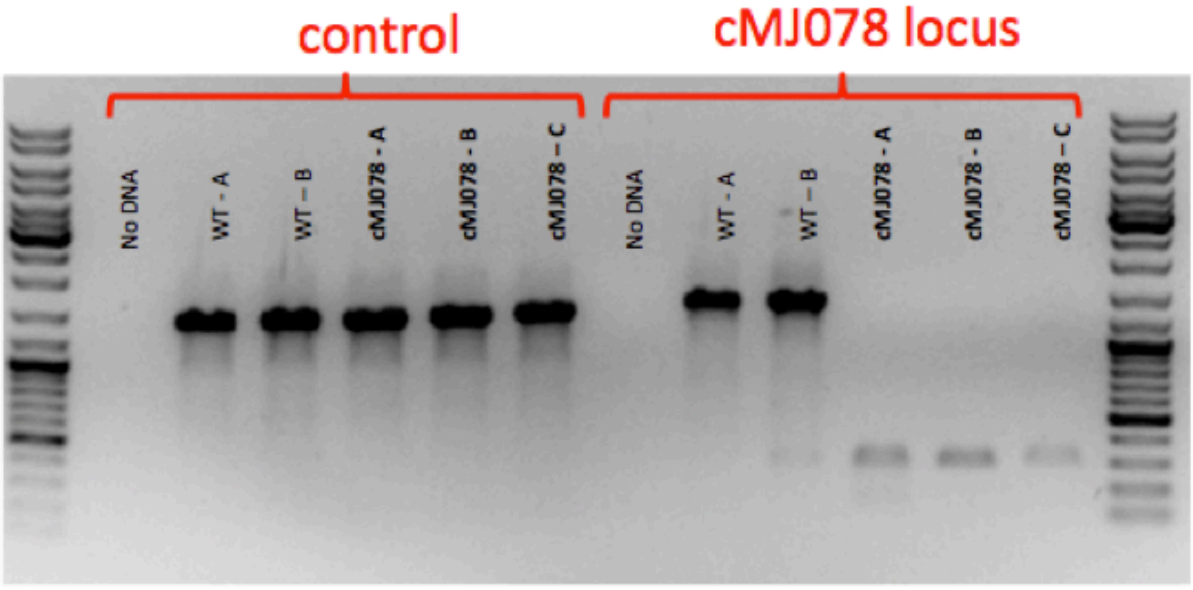

Figure 2: PCR indicates that all three colonies of CMJ078 have a disruption at the expected locus. Both WT and CMJ078 produce control locus amplicons of the expected size (1315 bp). For the PCRs across the locus disrupted in CMJ078, the wild-type colonies show a band of the correct size (1400 bp), but no band is present in the cMJ078 mutant strain.

IV. Amplify the cassette-genome junctions

Amplification of cassette-genome junctions is complicated by the fact that the cassette may be truncated on one side. Therefore, it is easiest to first check the side with the already-mapped flanking sequence, because the primer binding site on that side of the cassette is likely intact. You must be mindful of the possible cassette orientation in the genome for choosing the correct genomic-cassette primer combination. An example for a mutant generated with pMJ013b (same as LMJ.SG0182 mutants) is provided below. Note that for LMJ.RY0402 mutants, primers oMJ005 and oMJ155 should be replaced with oMJ913 and oMJ944 (see page 4 for sequences) respectively.

Example: For our samples, we used two sets of primers to separately amplify the 5' (primers oMJ005 and oMJ274) and the 3' (primers oMJ155 and oMJ276) cassette-genome junctions. cMJ078 is a mutant that contains an intact cassette that has inserted cleanly into the genome without cassette truncation (Figure 3).

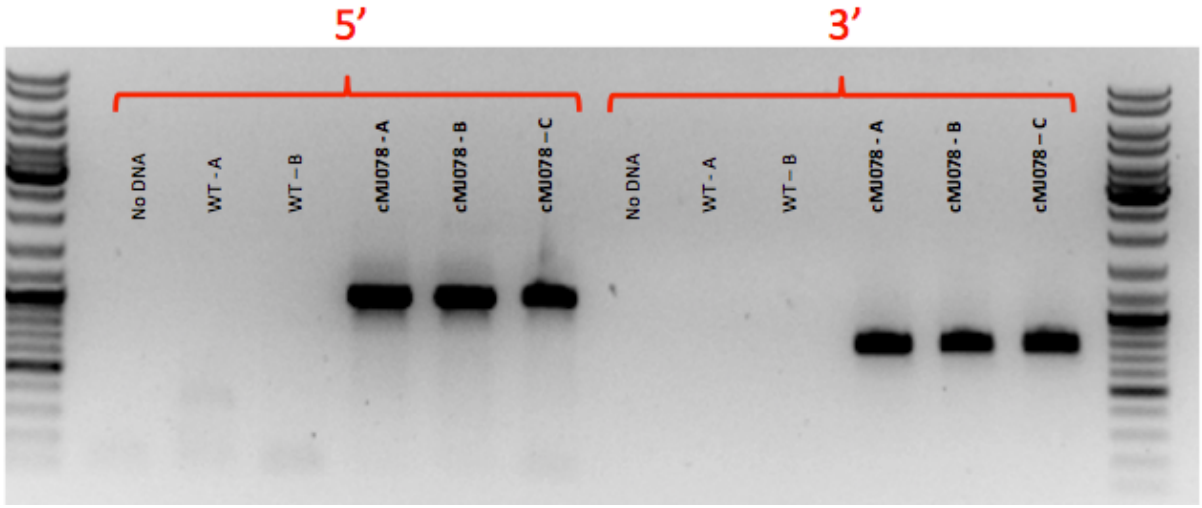

Figure 3. Genome-cassette junctions can be amplified from mutant strain cMJ078. As expected, the wild-type shows no product for either cassette junction, while the cMJ078 colonies have a product of the correct size for both cassette junctions (5': 980 bp; 3': 726 bp).

If the cassette is truncated on one side, it may be challenging to characterize the cassette-genome junction on the truncated side. Several primers binding to the middle of the cassette can be used in combination with genomic primers to amplify that side. Any amplification products obtained should be sequenced to determine the sequence of the junction. The sequence of the transformation cassette generated from pMJ013b can be found here:

https://www.chlamylibrary.org/insertion_cassette/pMJ013b-MlyI

The sequence of the transformation cassette CIB1 can be found here:

https://www.chlamylibrary.org/content/CIB1