

Validation of recombinant plasmids by Sanger Sequencing

NOTES: There are two distinct approaches to this, either sequencing the plasmid directly, or amplifying the insert region, and sequencing that. The advantage of sequencing the plasmid directly is that there is no chance of PCR mutations influencing the sequence, but the disadvantage is that we tend to get worse sequencing reads from plasmids compared to PCR products.

In either case, PURITY of the DNA is critical for success of sequencing. A good strategy is to first attempt to sequence the plasmid directly, and if that fails, then amplify the insert region ('spanning PCR' above), then sequence that. In the case of PCR products, these must be purified (e.g. via silica spin columns) *before* sequencing.

SEQUENCING PRIMERS: The usual primers to use for sequencing would be the F1-R2 primers in the diagram above. However, for large inserts (>1 kb), you will need to design and order additional primers, since the average sequencing read length is only 500-700 bp, depending on template quality. In that case, space additional primers about 400 bp from the F1 and R2 primers, to ensure you get coverage of the entire insert sequence. (and for very large inserts, place further primers again ~400 bp away from the second set of primers). For maximum confidence in the sequence, it is good practice to sequence BOTH STRANDS of the DNA, ie. to ensure you have duplicated sequence data all the way along the insert DNA.

Note that each sequencing reaction only contains ONE PRIMER! It is not a PCR!
See below for detailed instructions on setup of Sanger sequencing reactions for AGRF.

Preparing and sending samples for Sanger sequencing (AGRF)

General notes: this sequencing method is appropriate for PCR products or plasmids that are expected to be a single sequence. This method is NOT appropriate for diversity analysis (e.g. sequencing 16S PCR products from soil) or other samples where a mix of sequences are expected. A good Sanger sequencing run should yield at least 800 bp of sequence. If you are sequencing a large construct eg. a plasmid, you should space your sequencing primers about 500 bp apart to ensure overlap of sequence reads, even if the sequence is a bit shorter than usual. The first ~50 bp of sequence is junk, so keep this in mind when thinking about where to design primers.

Specific instructions:

Download the pdf "Sanger Sequencing: Sample Preparation Guide" from the AGRF website and read the relevant bits of this document. Especially note the table on page 5. Our samples are almost always the "PD" type (purified DNA).

Set up sequencing mixtures in 1.5 ml tubes with a short informative label on the lid. (e.g. NC1, NC2, NC3...). Wrap top in a little Parafilm to stop it popping open during postage.

Each mixture should contain:

- Column-purified PCR product: typically 1-2 uL (10-75 ng)
- Sequencing primer*: 2 uL of 5 uM stock solution
- Sterile MilliQ water: Remainder of mixture, to total of 12 uL

* when sequencing a PCR product, we usually use one of the initial PCR primers as a sequencing primer too but note that the sequencing primer solution is 10x more dilute. Use sterile MQ to make the dilution. Note that a sequencing reaction only needs ONE primer, not two. It is not a PCR.

Table 1: Recommended amounts of template and primer for sequencing reactions

Template	Recommended Quantity for <u>PD</u> Samples (in 12µL)	Recommended Quantity for <u>CS</u> Samples (BDT set up in your own lab)
PCR Product 100 – 200 bp	3 - 8 ng	1 - 3 ng
PCR Product 200 – 400 bp	6 - 12 ng	2 - 4 ng
PCR Product 400 – 600 bp	12 - 18 ng	4 - 6 ng
PCR Product 600 – 800 bp	18 - 30 ng	6 - 10 ng
PCR Product >800 bp	30 - 75 ng	10 - 25 ng
Plasmid, Single-stranded	150 - 300 ng	50 - 100 ng
Plasmid, Double-stranded	600 - 1500 ng	200 - 500 ng
Primer Quantity (one primer per reaction)	10pmol* (0.8 pmol/µl)	3.3 pmol

**this equates to 1µL of a 10µM stock, or 2µL of a 5µM stock*

Please note that these template amounts are guides only and optimisation may be needed

Most plasmids we work with are only a few kb and so a couple of hundred nanograms is likely to give a good sequencing run. Making up your sequencing reaction with 10µL of plasmid prep to get 1500ng is more likely to give a bad sequencing reaction than 200 ng with MilliQ added to dilute impurities. If you are working with a low copy plasmid, you may need to amplify sections of the plasmid by PCR and sequencing the PCR products rather than sequencing the plasmid directly.

Assemble your sequencing mixture on the same day as you plan to send it, ideally just before posting it. Even purified DNA will degrade eventually at room temperature, so we want to minimise this exposure time. Wrap your tubes in bubble wrap or other protective packaging, and add them to an express post envelope (available in G08 store).

Next you need to fill out the online sample submission form on the AGRF website. (www.agrf.org.au) Get the login name and password from Nick.C.

Go to the 'submit samples' tab, then choose service type 'Sanger sequencing'. Region is 'Sydney', service is 'Purified DNA', sample format is '1.5 ml tube', then choose the correct product type for your samples, then base call method should be 'N calls for low-quality bases'

Fill out the table at the bottom of the page. It's very important that the sample names entered in this table are *exactly* the same as what is written on your tubes. (You don't need to worry about the 'sample name range' or 'sample prefix' boxes.) Don't tick any of the safety boxes !

Choose purchase by 'Cost Centre' then type in your lab account code as the cost centre code (check with Nick.C. for the account code). The delivery method is 'Post' then in the next box choose 'Express Post'. Temperature is 'Ambient' Check everything is entered correctly then hit Submit. **You will get a pdf version of the document, which you need to print and add into your package.**

Seal the express post envelope and address to:

Australian Genome Research Facility Ltd.
Westmead Millennium Institute Building
176 Hawkesbury Road
Westmead NSW 2145

Make sure you sign the hazardous good declaration on the back and peel off the tracking sticker so you can track the package. Post in a yellow express post box. There is one of these on Science Rd (outside the uni post office) and another on King Street (Newtown), near corner of Missenden Rd.

Nick gets an email when your sequencing results are ready (usually 2-3 days after they receive the package), then I will let you know. You can then log in to the AGRF site and retrieve your data. Go to Retrieve Data tab, then Click to connect to AGRF cloud (you may need to enable popups)

The sequence data comes back as several different file types. The .fa and .seq files have basically the same information (i.e. your sequence) ; these can be opened in any text editor (eg. Word or Notepad) and will also open with SnapGene. The .ab1 file has the raw data from the sequencer. If your data is good, you probably wont need the .ab1 file, but if you have N's in the middle of your data you may need to check the raw data and correct the sequence based on manual inspection of the locations which have given 'N's. The ab1 file will open with SnapGene or other dedicated sequence viewer programs. Note that ALL Sanger sequence data will have N's at the start and at the end, these should be removed before doing any further analysis.