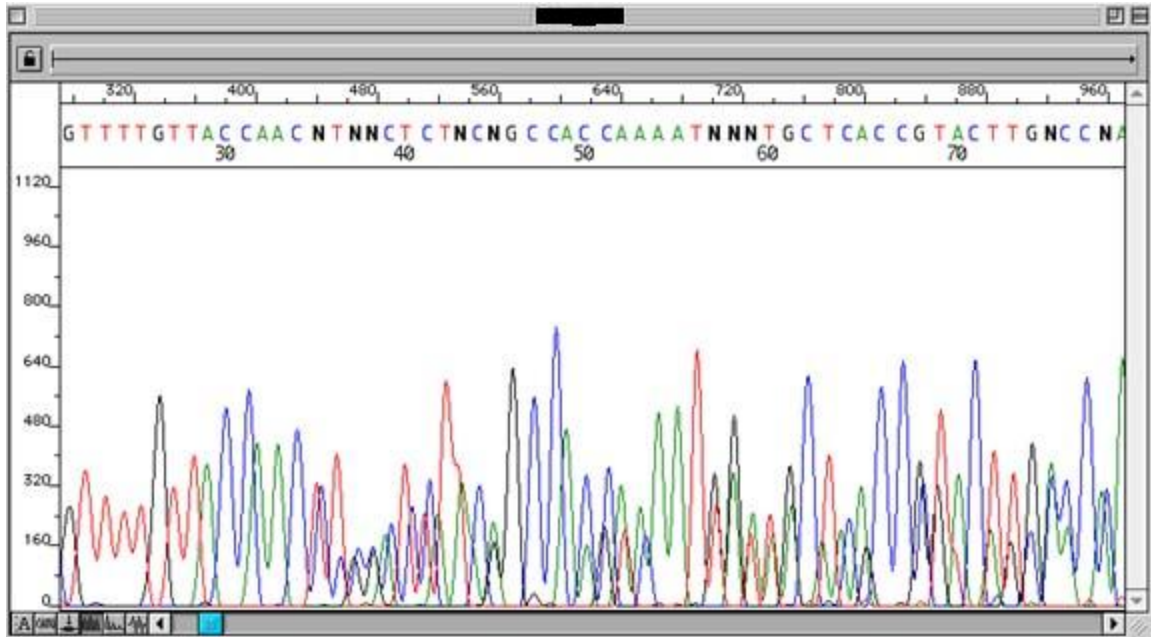


DNA Sequencing: Troubleshooting

Below you will find some common problems associated with DNA sequencing as well as the possible causes and solutions for these problems. A picture of the sequence traces is presented above the information describing the problem, how to identify the problem, the cause, and the potential solution for the problem. Other problems can occur with sequence data, but the ones presented below are those seen most commonly.

Compression Sequence



Problem: Sequence appears as multiple overlying traces, not always throughout the entire sequence

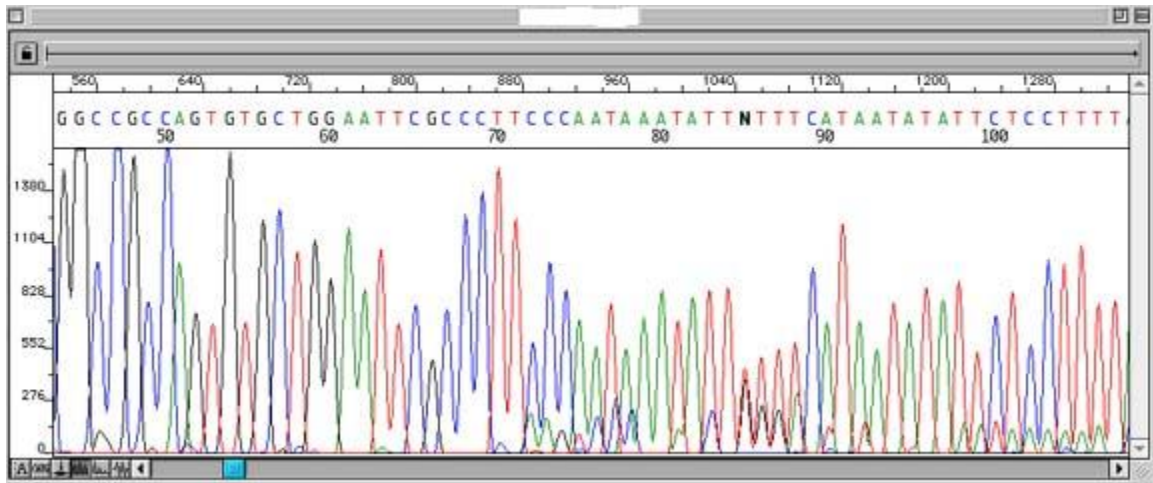
How to Identify: Sequence data begins to show multiple overlapping traces after a point in the sequence, although the sequencing signal remains strong. A compression phenomenon and contaminated DNA prep can look similar, though the contaminated prep will typically show overlapping traces immediately after the insert cloning site.

Cause: DNA fragments of different size with the same electrophoretic mobility, ie fragments that migrate on top of each other during electrophoresis. This phenomenon is thought to be caused by regions of secondary structure within the template DNA, which are often, but not always, found in regions with a high G/C or high A/T content.

Solution:

1. Add DMSO to a final 5% (v/v) concentration to the sequencing reaction.
2. Linearize plasmids with a restriction enzyme.
3. If the sample is a PCR product, try amplifying the DNA with substitution of 7-deaza-dGTP for 75% of the dGTP in the PCR, and then sequence the PCR product.
4. Choose a new primer close to the compression site which can help avoid the effects of the secondary structure.

Contaminated Plasmid DNA Preparation



Problem: Contaminated plasmid DNA preparation leads to overlapping peaks in the sequence data.

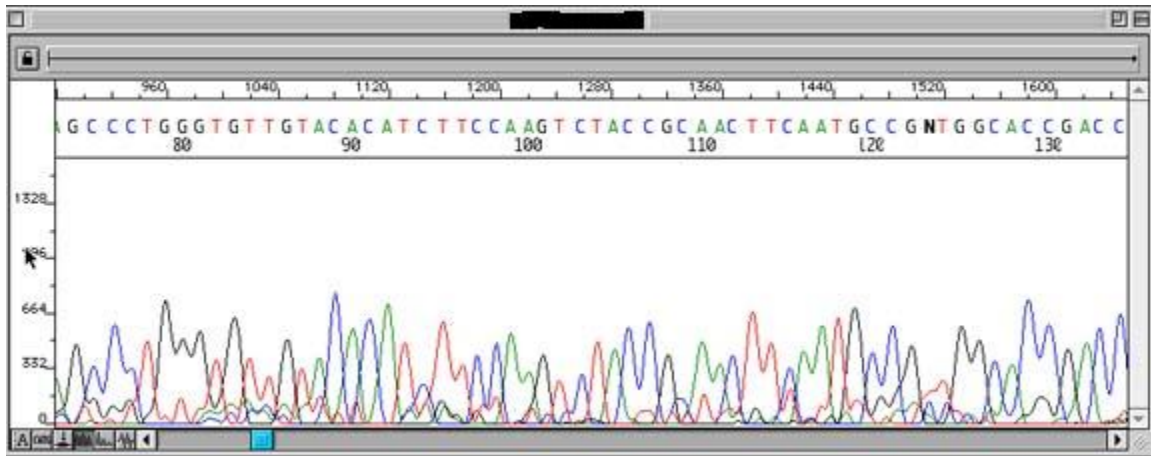
How To Identify: Sequence data is clear out to the end of the insert cloning site, at which point peaks underneath other peaks can often be seen. In the worst case, the peaks will obscure each other, making the entire sequence unreadable.

Cause: Picking more than one clone for a single plasmid preparation, or picking a single colony that contains two or more plasmids, each with different inserts.

Solution:

1. Carefully pick a single colony from your plate into your growth medium in a clean environment.
2. Verify that your single colony contains your insert of interest after DNA preparation by restriction digest and agarose gel analysis.

High Background/Noisy Sequencing Signal



Problem: Low/Poor Signal-to-noise ratio in sequence data causing unreadable sequence data at points in the sequence.

How to Identify: Observation of small peaks underneath all peaks in a sequence read. Low signal can be confirmed by checking the 'information view' (click on the 'A:' radio key seen at the bottom left of an EditView chromatogram trace) and verifying that Average Signal Intensity is less than 100 relative fluorescent units (rfu). In this picture, the average sequencing signal for each base was as follows: G (44), A (50), T (20), C (47). Note: It is also possible that 'double peaks' throughout a sequence are suggestive of the primer binding to more than one site on the DNA template. However, the fluorescent intensity numbers would suggest that the above results were caused by low signal.

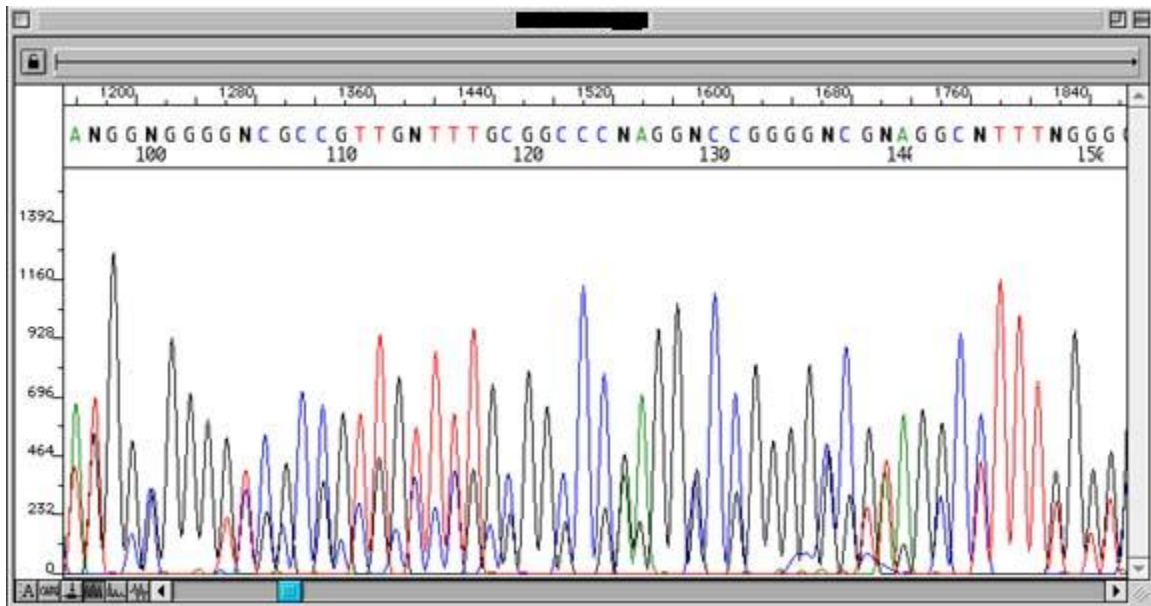
Cause:

1. Too little DNA template is added to sequencing reaction.
2. Primer binding to the template is not very efficient.
3. The sequencing reaction was not very efficient, as is sometimes the case with BAC end- or direct BAC sequencing.
4. DNA template was not purified well enough prior to the sequencing reaction and is contaminated.

Solution:

1. For (1), increase the amount of template added to the reaction.
2. For (2), increase the amount of primer in the reaction, or if possible, design a new primer.
3. For (4), you can shear BAC DNA by vortexing it at maximum rpm for 1 minute or passing it through a 28 gauge sterile needle, digest it with a restriction enzyme prior to sequencing, and denaturing the BAC DNA prior to adding to the sequencing reaction (98C for 10 min).
4. Prepare the DNA again using fresh reagents and careful techniques.

N-1, N-2, N-3... Priming Sequencing Trace



Problem: Sequence shows 'double peaks' throughout, with the second (generally smaller) peak being the same base as that of the true base immediately to the right of it.

How to identify: Although an n-1 primer problem may look like compression or priming of two different regions of DNA template, it can be readily identified by each position containing the true peak and the peak immediately to the right of it.

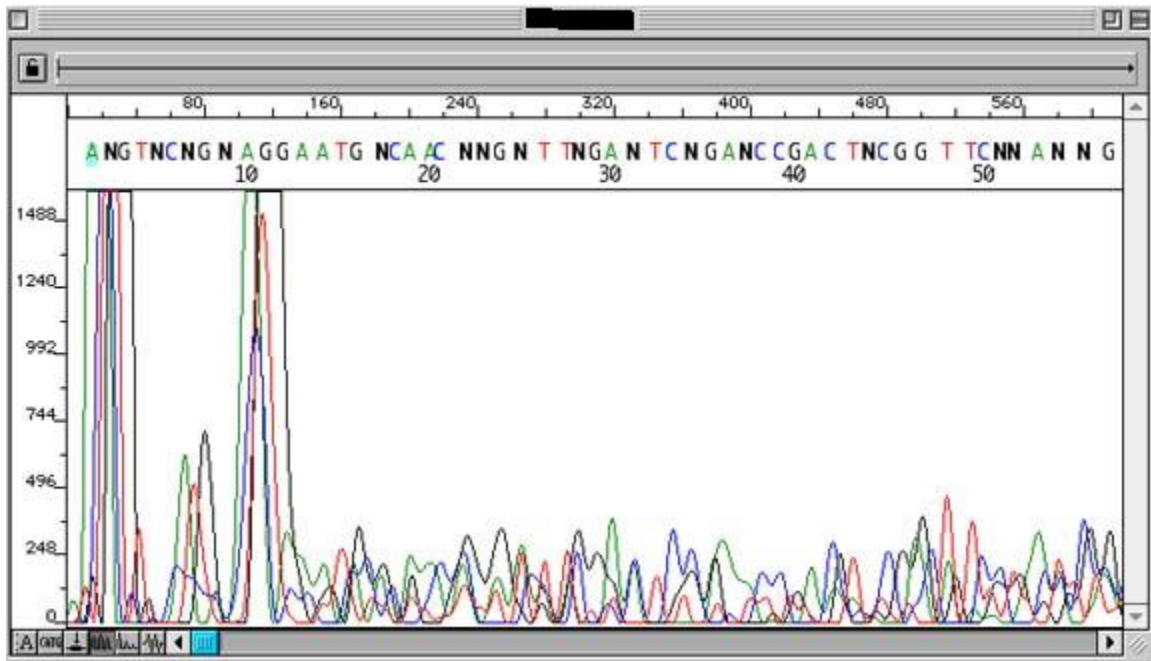
Cause:

1. Poor purification of the oligonucleotide during primer synthesis which leaves a certain percentage of n-1mers in the final product. When the DNA template is sequenced, this percentage of n-1mers will prime the DNA template, causing some of the sequence to be 1 bp shorter than it should be. Primers that have begun to degrade will also do so from the 3' end, causing a proportion of the original sequencing primer to become n-1.
2. The primer has started to degrade from the 3' end so that a portion of the sequences will have a sequencing ladder that is one basepair longer than a non-degraded primer.
3. Poor primer binding sight on the template.

Solution:

1. Cartridge purification of oligonucleotides and even HPLC purification will help insure that n-1 priming of sequence doesn't occur. In both cases, re-synthesis of the oligonucleotide is necessary.

No Readable Sequence Data



Problem: Lack of sequence data

How to identify: Sequencing trace is non-uniform. Unincorporated Dye Terminator Peaks at the beginning of the sequence are out-of-scale. Basecalls contain many ambiguous bases ('N').

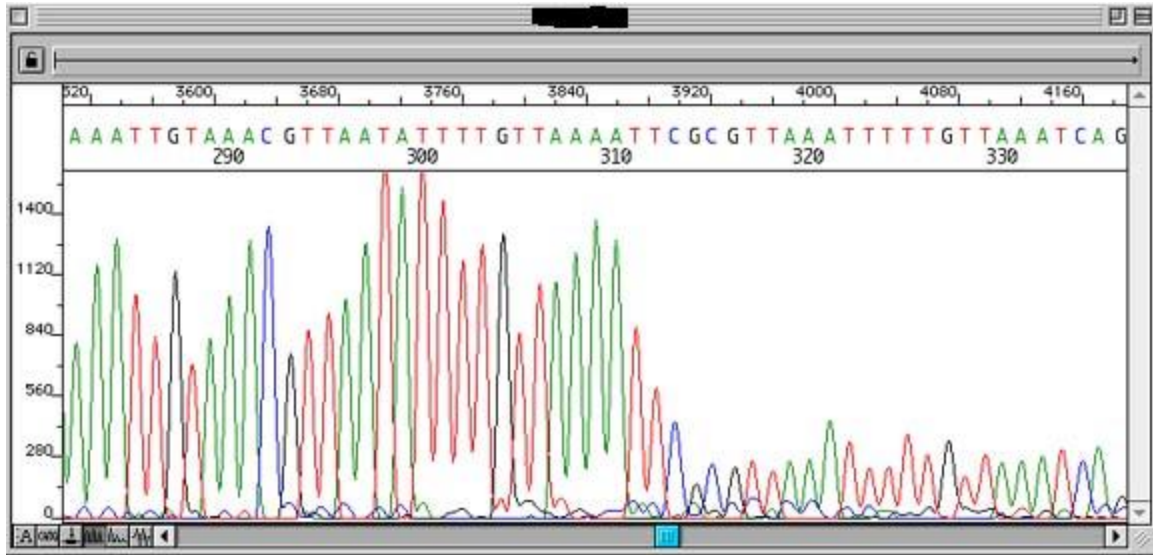
Cause:

1. Primer not annealing to template (no priming site).
2. No DNA template present in sequencing reaction.

Solution:

1. For (1), design a new primer which is free of hairpins, free of duplexes, and has a T_m greater than 55°C.
2. With (2), it is important to verify DNA concentrations before addition to the sequencing reaction in order to insure that the optimum amount of template is being added to the reaction. It is best to verify your concentrations visually by running them on an agarose gel.

Secondary Structure



Problem: Sharp drop or premature termination of sequencing signal during the sequencing run.

How to identify: Sequencing signal drops off sharply or terminates before the end of the sequencing run (termination of sequencing signal also occurs with PCR products, depending on how long the PCR fragment is).

Cause: Secondary structure in the template can cause various anomalies that result in inefficient chain elongation after the region of secondary structure.

Solution:

1. Many of the same techniques used to eliminate compression can be helpful in obtaining good sequence data past regions of secondary structure. Although not always practical, changing sequencing chemistries can produce better results as well.