# How Do I Read + Crop Sequences?

Reading them:

1. Is it readable or not readable?
   a. Look for clear, distinct peaks.
   b. Look for any "Ns" in the top-this marks an error in sequencing (the machine is unsure what base it is). If there are Ns but you are certain the machine misread, you can edit out the N and put in the correct base.
   c. Check for the landmarks: the starting sequence CGGCCGGG(GG...) and the poly-A tail. <mark>There are certain cases in which you do not have a poly-A tail but readable.</mark> See further below if this is your case.
   d. If you meet all the requirements, **you have a readable sequence!** Now you can start cropping.

No "Ns". If you know the correct base and the machine misread, you can edit the Ns to the correct base.
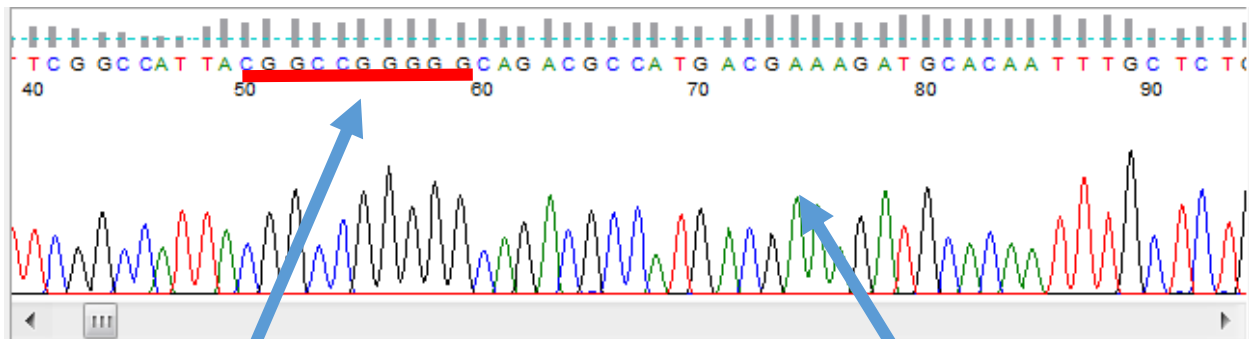
The first "landmark", CGGCCGGGGG marks the beginning of the sequence. The first base is **C60**.

Peaks are clear and distinct.

Distinct peaks and no "Ns".

AT GA GG T T TT TT AA A A AA A A AA AA AA A A A A AA A A A A AA A A A A
810   820   830   840

Poly-A Tail. The sequence ends on the last base of this tail.

2. Unreadable situations: What now?
   a. Even if your sequence is unreadable, you should analyze the reasons why your sequence is unreadable to prevent future mistakes. These are the three most common situations, the problems that caused it, and how to prevent it.

No. 1: The "Blank Sequence":

EX3–NS1.ab1

4Peaks   Search

5 bases – Nothing selected

1.0x

The above picture is a completely empty sequence. This is caused by two mistakes during the lab procedure.

1. You messed up removing the SDS, or the detergent.

The detergent is responsible for destroying the bacterial cells and everything in it. This helps us extract the plasmid DNA. The plasmid DNA recombines afterwards, letting us sort it out from the rest of the mess. However, if we leave the SDS in the mix, the plasmid DNA will not recombine. Your sample will come up as empty.

2. You messed up removing the ethanol.

The Short Answer: Ethanol removes DNA. We use it to get rid of contaminate fragments as well as nucleases, an enzyme we use. If we keep ethanol in the mix, plasmid DNA will "fall out" of the mix as well, and it will be empty.

The Long Answer (You can skip if not interested, but it's very good to know, esp. for anyone interested in working in biochem labs):

DNA is a polar molecule, negatively charged due to its phosphate groups. Ethanol, an alcohol, is also negatively charged. Due to the high polarity of water molecules, the electrostatic forces between water and DNA molecules are much lower (Coulomb's Law!) and the DNA easily dissolves. In fact, water molecules form a hydration shell around the charged DNA molecules, weakening ionic bonds and dissolving the DNA.

Adding ethanol "distracts" the water. Ethanol is much less polar than water, and adding it will disrupt the hydration shell going on around the DNA. Ethanol, negatively charged like the DNA, will replace the water, and electrostatic forces go up, repelling the DNA. This method of DNA precipitation, as it's called, is only successful when adequate ethanol is added, generally around 60% of the solution. In the end, the DNA "falls out" of the solution due to increasingly stronger ionic bonds between DNA molecules and weaker ionic bonds between water as more Ethanol is added.

tldr; You left in some chemicals that ruined/got rid of your DNA.

#Jelisa
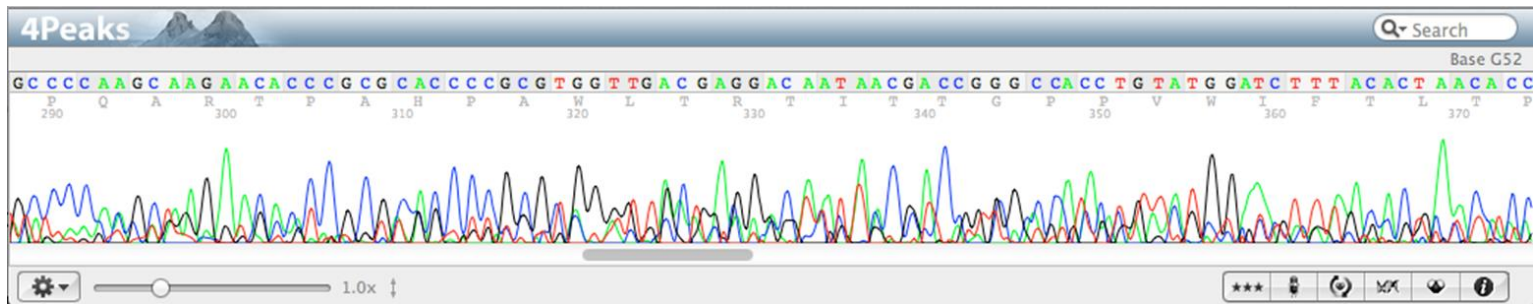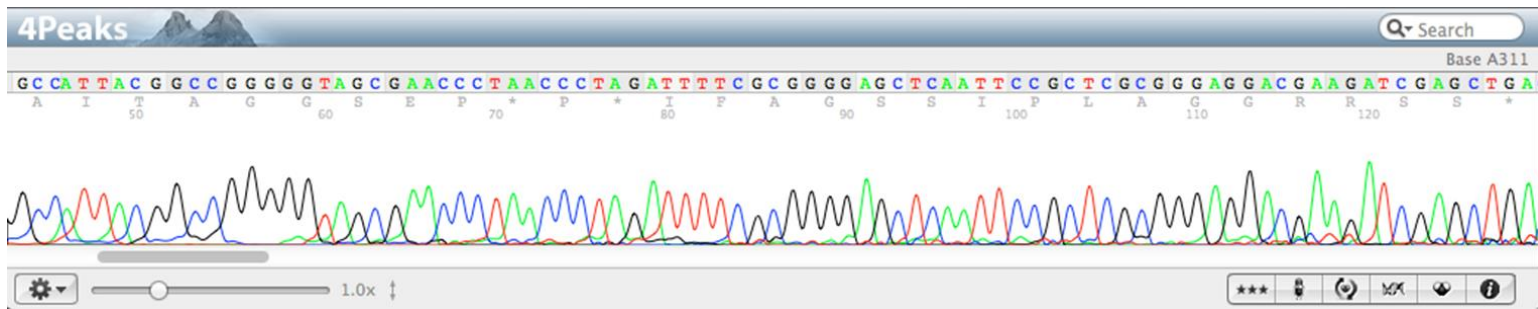
No. 2: Low Yield

In the above picture, the sequence starts out fine until it becomes super messy. Below, it gets crazy from the beginning. Both are classic cases of low yield.
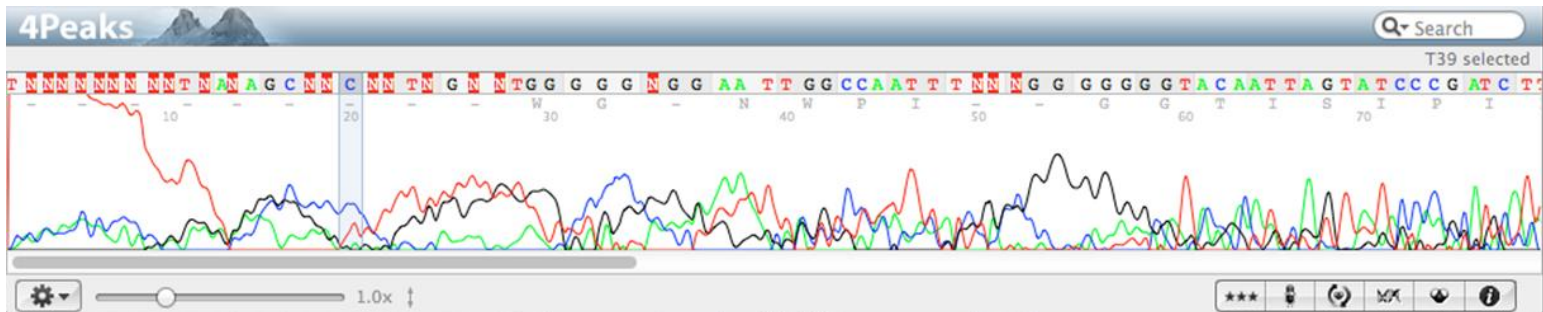

Low yield is caused by a small sample size. That means your bacterial colony did not grow to large enough quantities, causing far too little DNA sequences for the machine to read. There may be some other errors, such as pipetting badly and losing your sample size bit by bit. This should be a lot simpler than the previous one! ☺

#Jelisa

No. 3: The Double

C CAT TAC G G C C G G G G G GAT CC CAA C C C TAT C C C G A G
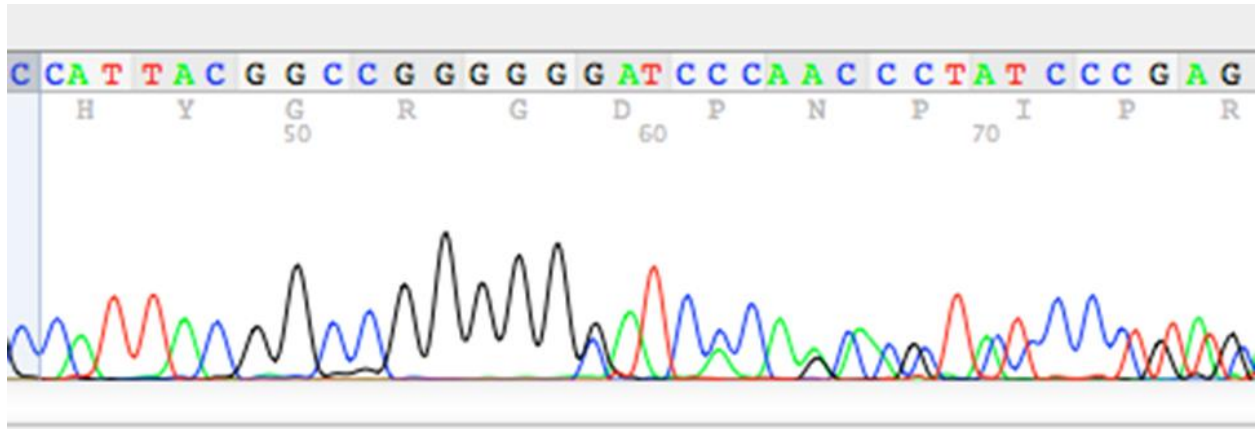
Above you can see that the vector backbone is very nice, but is messy when the DNA sequence starts. That's because you contaminated it with two (or more) sequences! See the illustration below for a clearer idea.

#1

#2

vector   insert

© 2014 WSSP
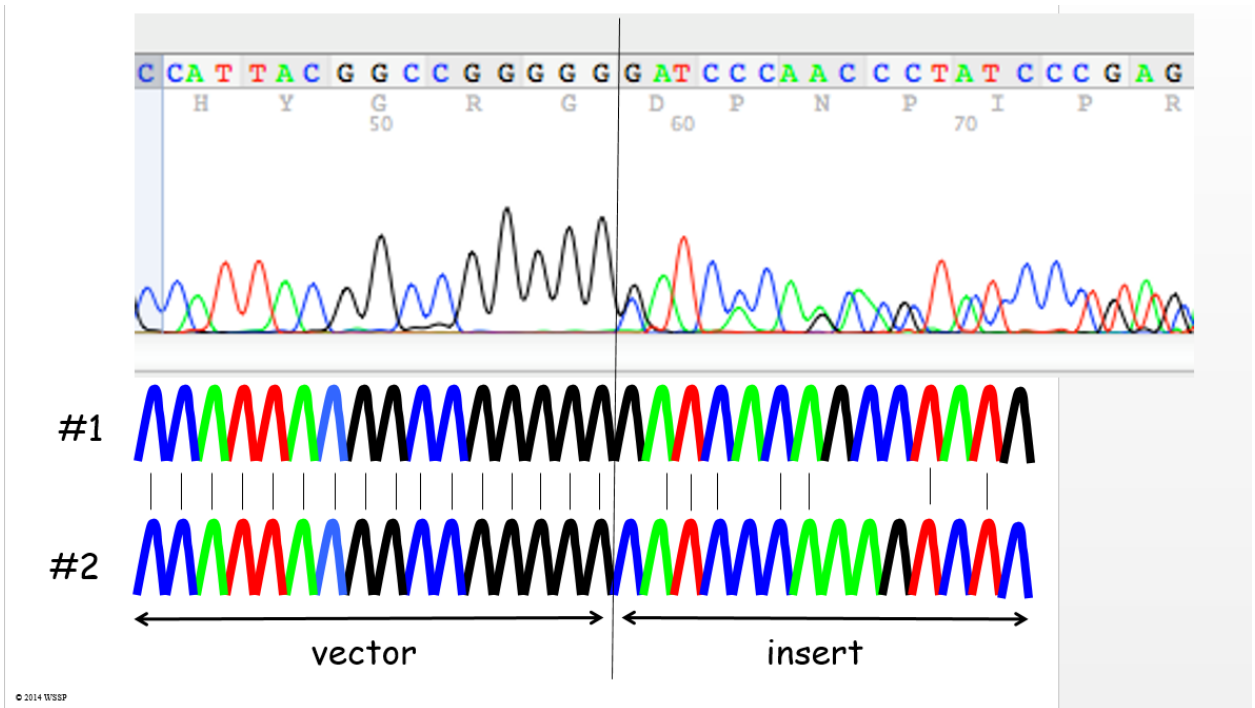
This can be due to the fact that you accidentally added someone else's DNA or picked 2 colonies by accident!
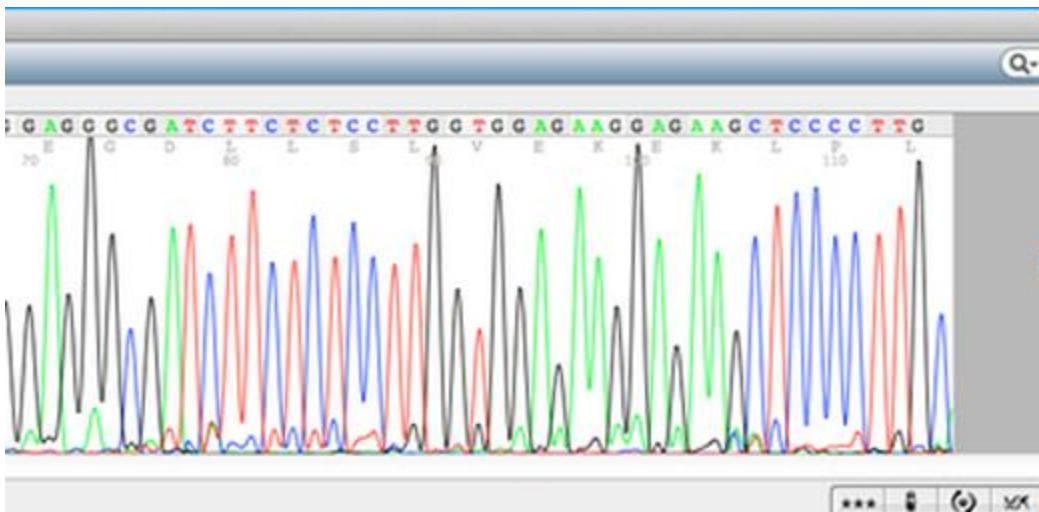
#Jelisa

3. What if there is no poly-A tail?

Simply proceed along with the DSAP analysis as per usual. This is not common, but you may still be able to find the corresponding protein in the Blast steps. You can still publish partial clones! Just be sure to mark it for future scientists.

4. Weird Cases.

These are exceptionally strange cases...We will go over them in brief so you will be exposed to them if the issue arises.
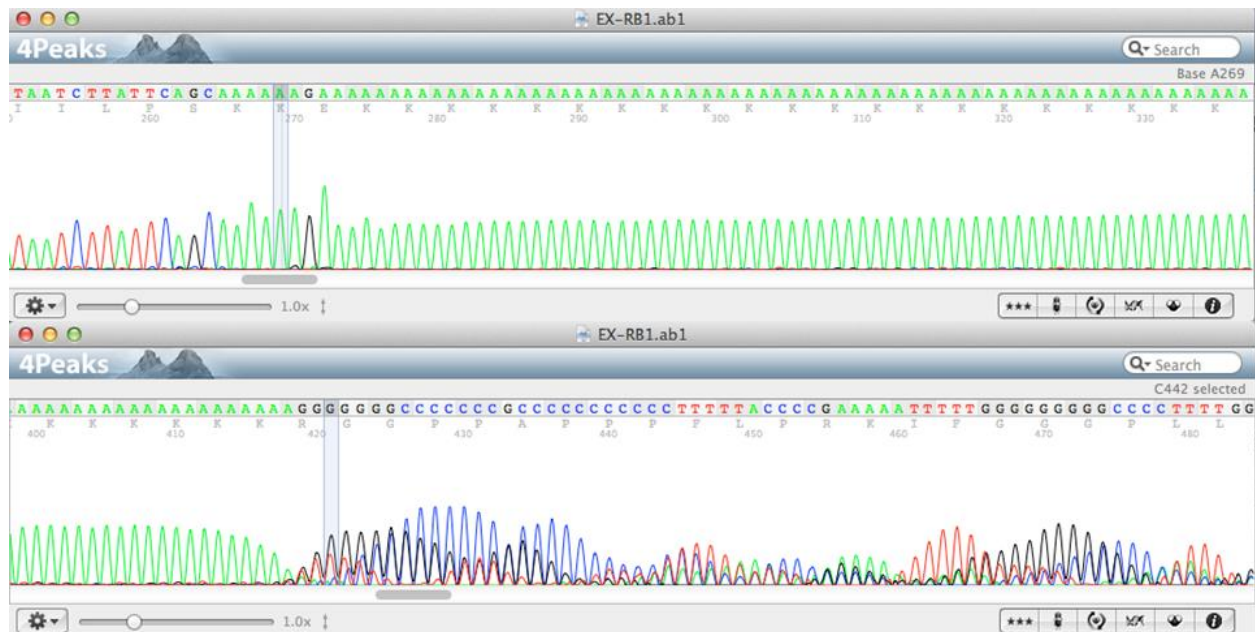
The Quick-Stop: The sequence is too short, so the machine just dropped it and stopped reading it.

This weird thing:

Sometimes, after a long string of a base pair, the machine starts to mess up. This then becomes unreadable.

The Wave:

The enzymes decided to add and take away some base pairs, resulting in this wave-format.

That's all! Have fun cropping! ☺