

3. Validating Results

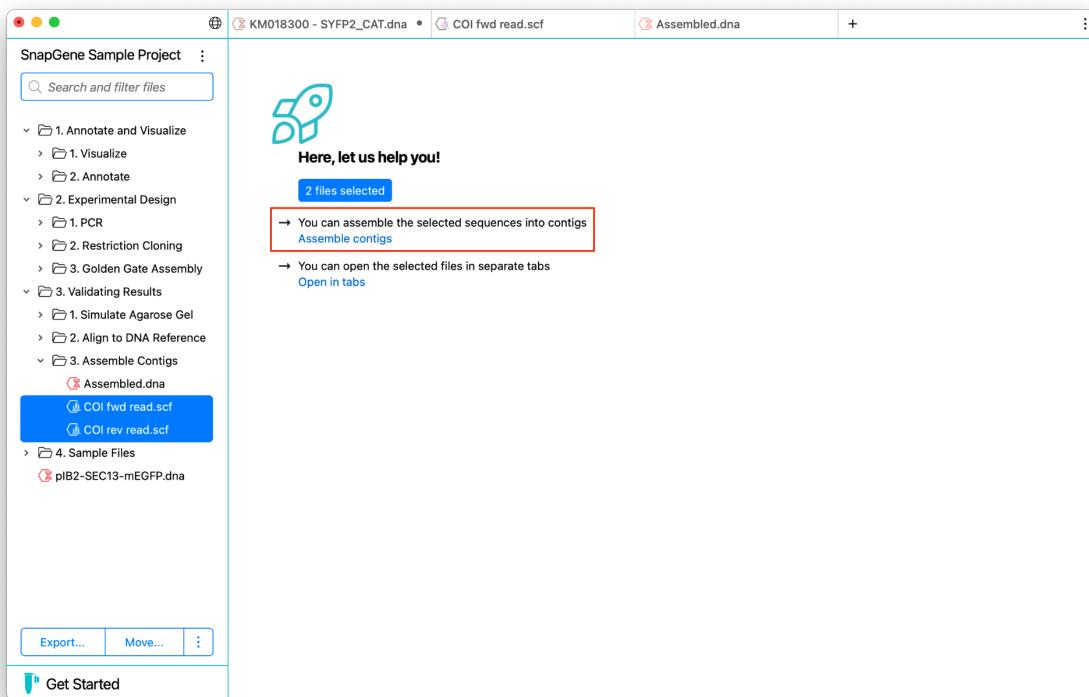
Assemble Contigs

The SnapGene **Assemble Contigs** tool uses the CAP3 assembler to assemble reads into one or more contiguous assemblies. This tool is designed primarily for assembly of a small set of Sanger reads, all derived from the same clonal source, and all of which are expected to overlap to form a contiguous sequence.

This folder contains 2 Sanger reads, one in the forward direction and one in the reverse,, corresponding to the cytochrome oxidase gene. The **Assembled.dna** file contains the reads assembled using the Assemble Contigs tool, while the remaining 2 chromatogram files (**COI fwd read.scf** and **COI rev read.scf**) are the forward and reverse reads that were assembled.

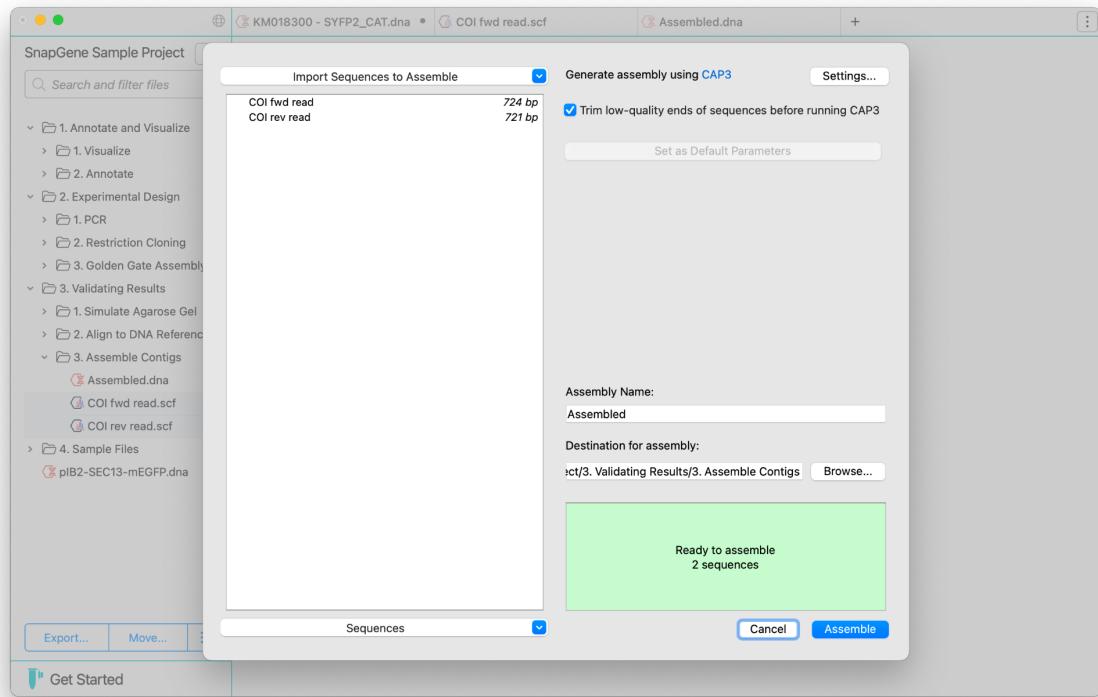
To replicate this assembly, do the following:

1. Select the two read files (**COI fwd read.scf** and **COI rev read.scf**). This will bring up the following prompt in SnapGene.



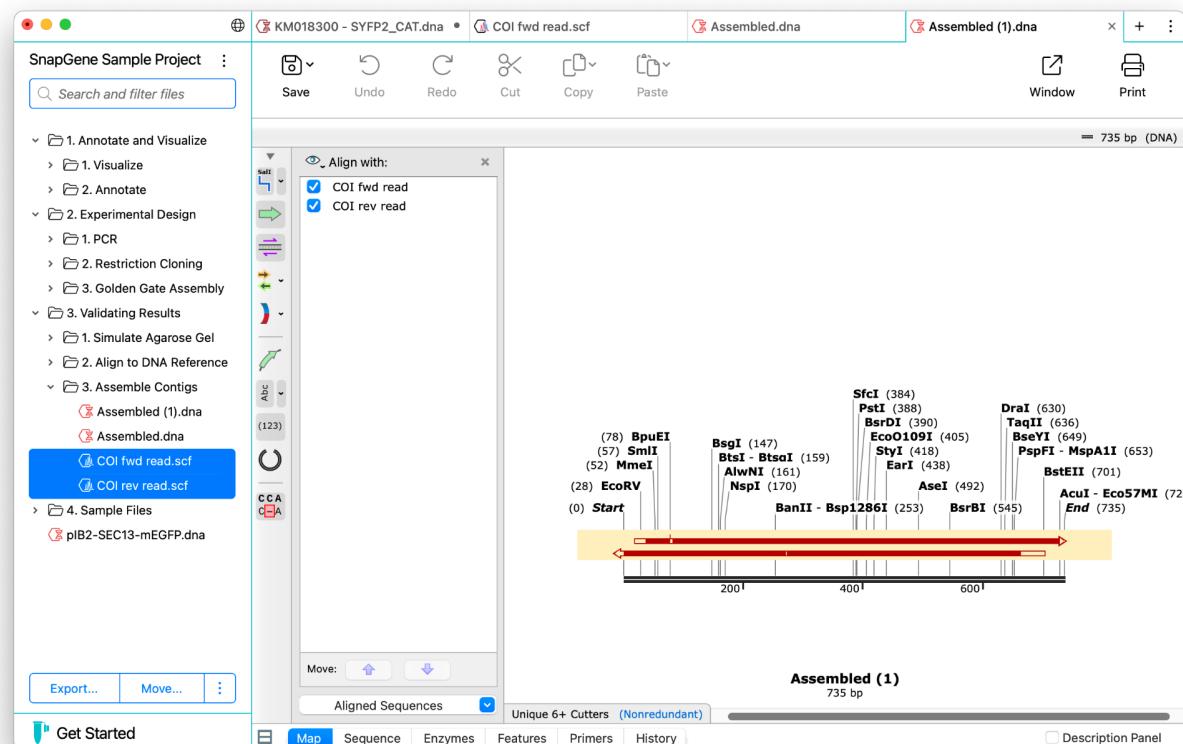
Select the **Assemble Contigs** link in the prompt. Assemble Contigs can also be accessed from the tools menu.

2. This will show the Assemble Contigs dialog. As you have selected your reads prior to opening this dialog, they will automatically populate the dialog. However, it is also possible to select the files for your assembly from within this dialog.

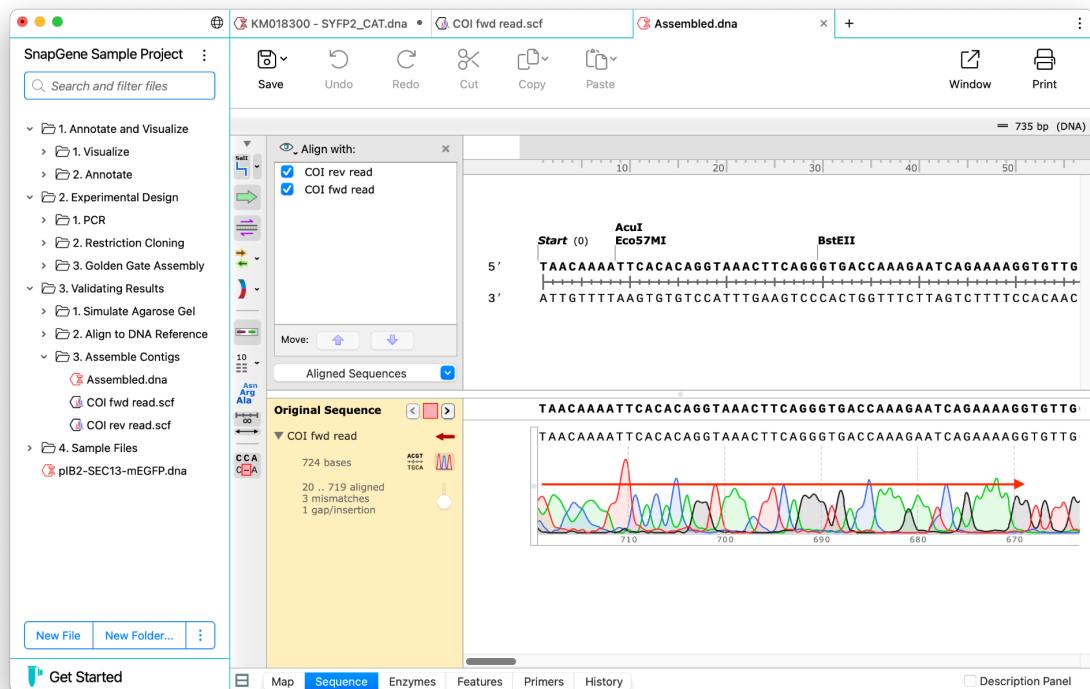


We will use the default settings for CAP3 for this assembly, but these can be adjusted by selecting the **Settings...** button on the top-right of the dialog. You can also opt to trim your reads prior to assembling them. Select the **Assemble** button to Assemble the reads.

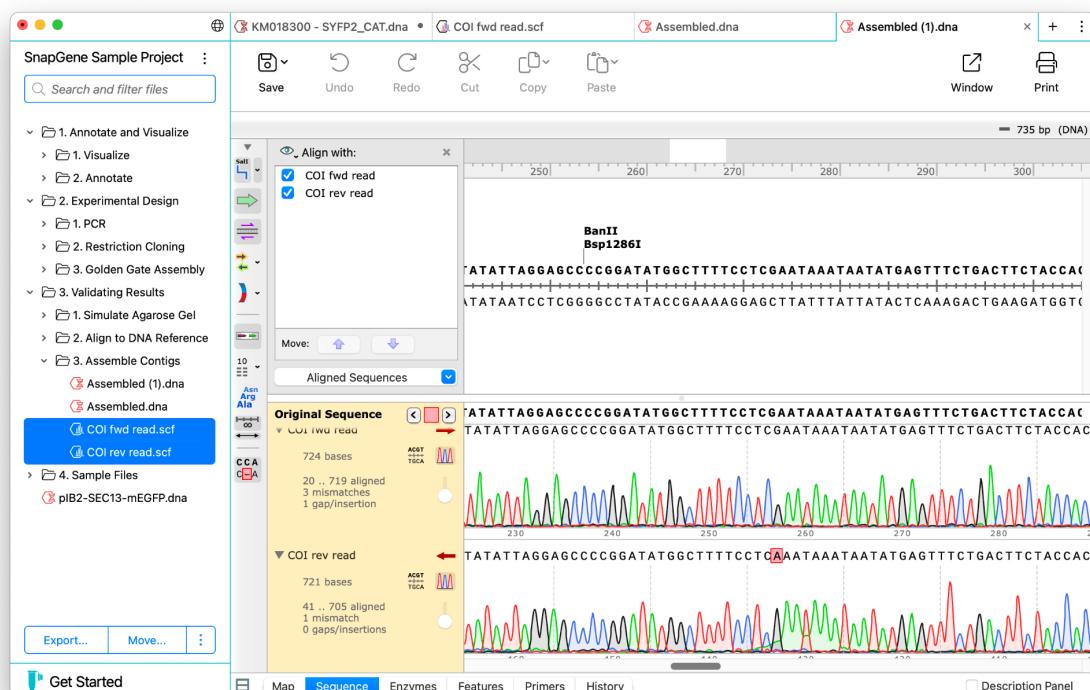
- This will bring up the Assembled.dna file in SnapGene. The Map view will show the overall assembly of the two sequences. There is a considerable amount of overlap between these two reads, but it looks like these reads might have some low-quality regions at the ends that do not match as well.



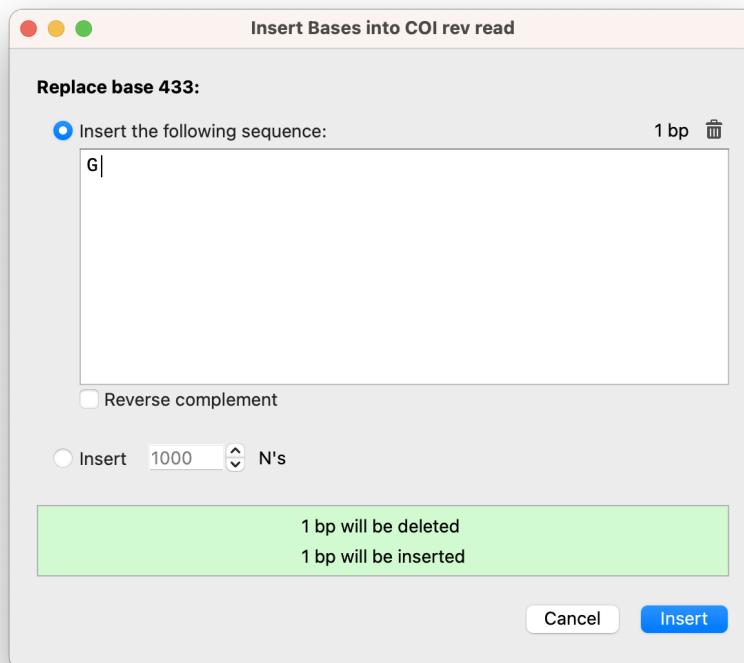
4. Switch to the Sequence tab by selecting the Sequence tab at the bottom of the window to see this assembly in more detail. If desired, the ends of the sequences can be adjusted by selecting the slider at the end of the sequence and dragging it to show or hide more of the sequence.



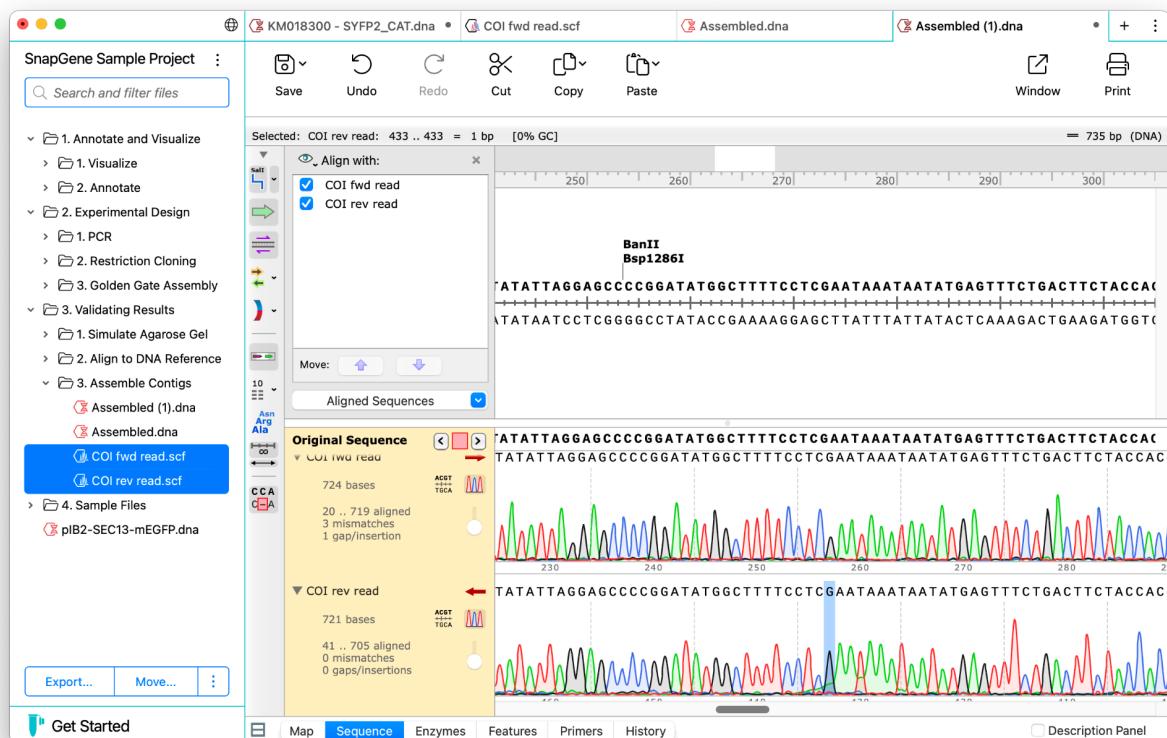
5. Scrolling along this assembly also reveals a mismatch between the two reads. Upon closer inspection it looks that this base might have been called incorrectly by the sequencing machine in the second sequence. A slight green peak corresponding to a G is present, and this has been called despite the far larger black (A) peak being present.



6. This can be edited by selecting the incorrectly called A base and selecting G on your keyboard. This will bring up the Insert bases dialog. Ensure that a G is being inserted then select the Insert button to insert this base.



This will correct the error in this sequence so that the two sequences now match perfectly over this region.



For more detailed information about the **Assemble Contigs** tool in SnapGene, refer to our [User Guide](#).