USING CONSED GRAPHICALLY

July 2014

Revised from http://bozeman.mbt.washington.edu/consed/distributions/README.20.0.txt by Emily C. Furbee, Amanda Boozalis, Rachel Greenstein, Michelle Itano, Andrew Nylander, and Wilson Leung

1) STARTING

Unlike most programs where you can open the files you want to work on within the program itself, you must launch Consed from the "edit_dir" subdirectory of the project you wish to work on. This tutorial will use a project named **Standard3** and we will assume this project is located in your home directory (the house icon in the Mac Finder, designated ~/ in UNIX).

Launch X11 and type the following in your xterm and then press **Enter** or **Return**:

```
cd ~/Standard3/edit_dir
```

Note that there is a **space** between "cd" and "~/Standard3/edit_dir". If the project is located in an alternate location, please consult with your instructor for the correct path to the **Standard3** project.

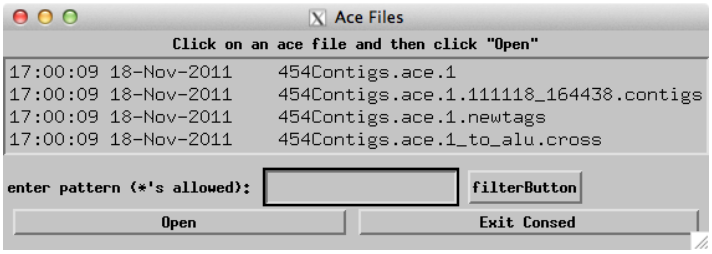
2) RUN CONSED

Start Consed by typing the following command in your xterm:

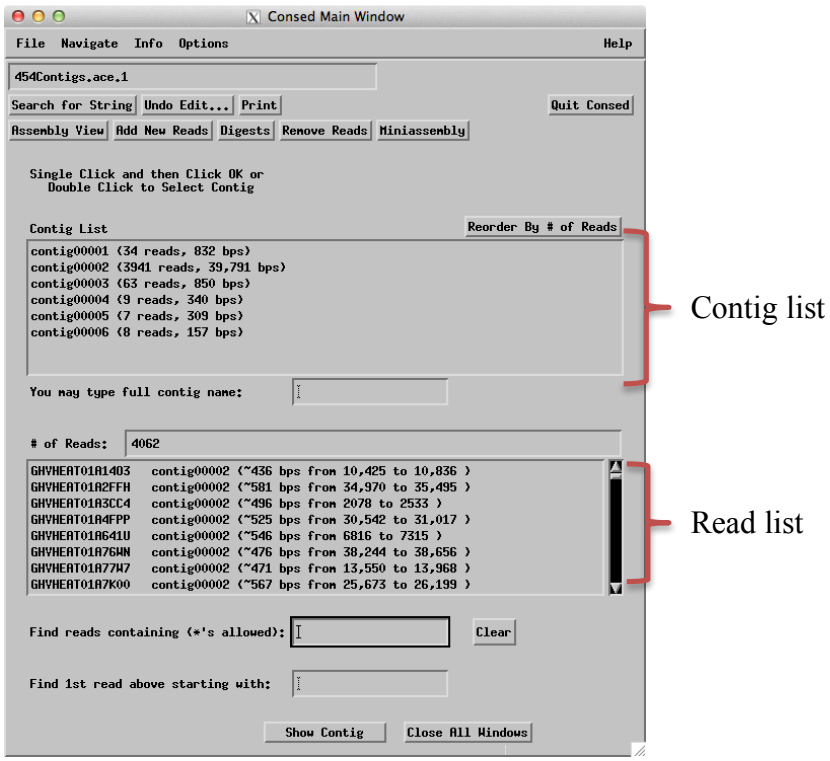
```
consed &
```

Two windows will appear. One of these will show the list of ace files in "edit_dir" and say 'Click on an ace file and then click Open'. Double click on "**454Contigs.ace.1**". .ace files are listed in chronological order with the most recently saved ace file listed first.

As you work on a project, Consed will keep track of the changes you have made to the assembly in an edit history file (.wrk file). If you open an ace file with a corresponding .wrk file, a dialog will appear which asks if you would like to apply the edits in the history file, you should answer "No" unless Consed crashes and you are trying to recover unsaved work.

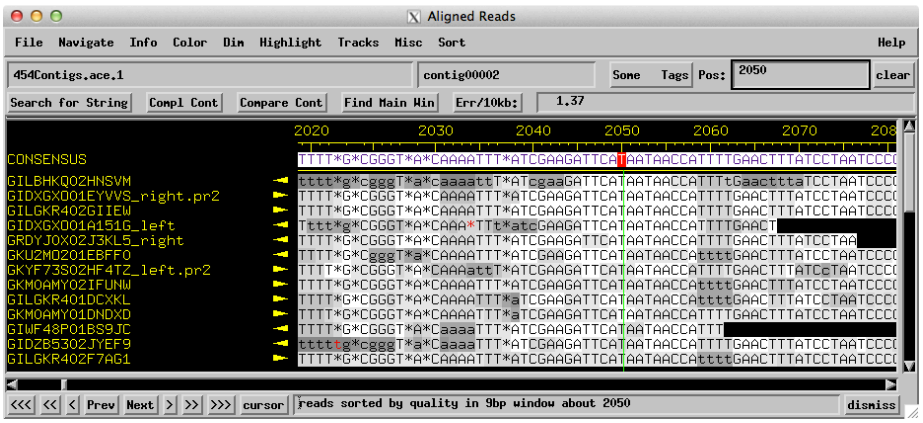


You should now see the 'Consed Main Window' with six contigs and each with a different number of reads. Contigs are listed in the Consed Main Window under "Contig List" from the smallest to the largest. If you double-click on one of the reads in the '# of Reads' section,' an 'Aligned Reads Window' will appear. Double-click on a read for 'contig00002'.



3) SCROLLING

Try scrolling back and forth. Try scrolling by dragging the thumb of the scrollbar near the bottom of the aligned reads window. Also try scrolling by clicking on one of the four '<<,' '<,' '>,' '>>' buttons for scrolling by small amounts. For scrolling by tiny amounts, click on the arrows at either end of the scrollbar. For scrolling by huge amounts, use the middle mouse button and just click on some location on the scrollbar. You can scroll to the beginning or the end of the contig by using the '<<<' or '>>>' buttons.

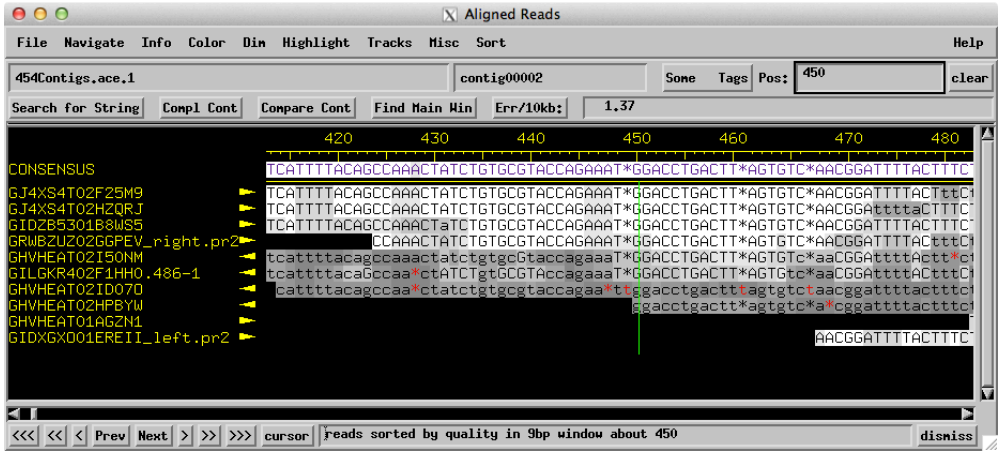


(Question: Why can't you just move the scrollbar to the extreme right in order to go to the end of the contig? Answer: In typical assemblies, there are reads that protrude beyond the beginning of the contig and reads that protrude beyond the end of the contig. Moving the

scrollbar to the extreme right will scroll to the end of the rightmost read – typically far to the right of the end of the contig. Thus, you typically use the '<<<' and '>>>' buttons.)

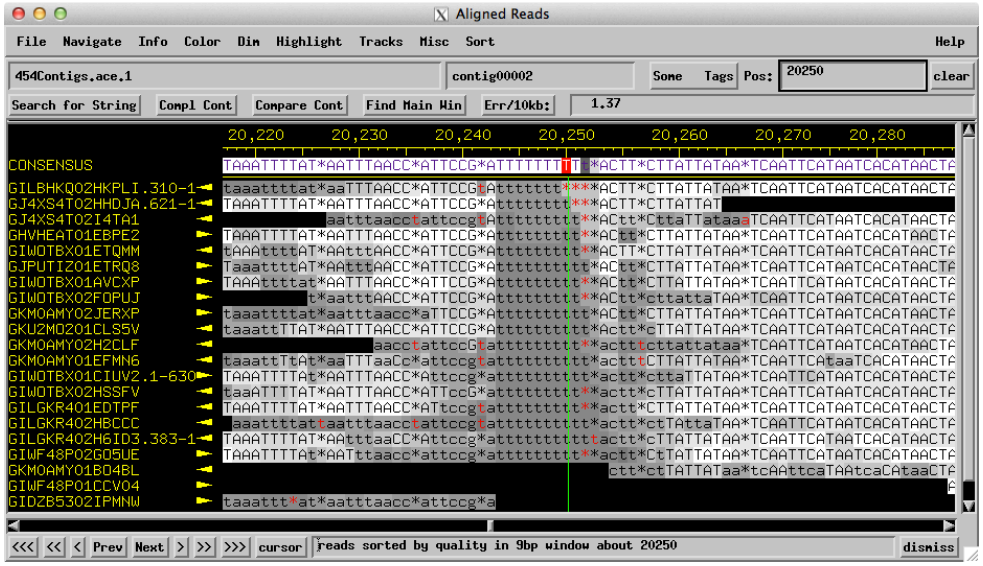
4) GO TO POSITION

In the Aligned Reads Window, click in the '**Pos:**' box in the upper right-hand corner. Type in a number, such as '450,' and push the '**Return**' or '**Enter**' key. The Aligned Reads Window will scroll to position: '450.' We find this feature is particularly useful when one person wants another person to look at something in the sequence.



5) COLORS

Notice the colors. Scroll to position '20250' and notice the red letters scattered about. The red bases are the ones that disagree with the consensus. Many of the red letters are the pads (*) found at the right end of the long stretch of T's. This is very typical for data generated by 454 machines, which have difficulty determining the correct length of mononucleotide runs.



Notice the different shades of grey background (around the bases). They indicate quality.
A quality value of 10 means 1 error in ten to the 1.0 power
A quality value of 20 means 1 error in ten to the 2.0 power
A quality value of 30 means 1 error in ten to the 3.0 power
A quality value of 40 means 1 error in ten to the 4.0 power

and for quality values in between:
A quality value of 25 means 1 error in ten to the 2.5 power

Also notice the upper and lowercase. This is just a cruder indication of the quality of the bases.

A quality value of 40 through 97 is given by white (the brightest shade), given the high quality threshold is set to 40 by default. Quality values between 40 and 10 are represented by a grey scale with lower quality bases being represented by increasingly darker grey color.

A quality value of 99 is reserved for bases that have been edited and the user is absolutely sure of the base (i.e. 'high quality edit').

A quality value of 98 is reserved for bases that have been edited and the user is not sure of the base (i.e. 'low quality edit').

To see the quality value of a particular base, click on it with the left mouse button. You will see the quality displayed in the Info Box at the bottom of the Aligned Reads Window.

The ends of the reads often show bases that are grey and have a black background. These are the low quality ends of the reads or the unaligned ends of reads, as determined by Consed. They generally have a quality value of 10 or less.

6) KEYBOARD USAGE

You can also scroll to the end of reads by using keyboard commands. Click on a base in a read. Then, hold down the control key and type 'a'. You will move to the beginning of the read. Hold down the control key and type 'e'. You will move to the end of the read. (Emacs users will recognize these commands.)

7) HIGHLIGHTING READ NAMES

In the Aligned Reads Window, click on a read name with the left mouse button. The name will turn magenta. Click again and it will turn yellow again. Try turning it magenta and then scrolling. This feature is helpful in keeping track of a particular read as you scroll.

If you have an Emacs window open (or any X11 window), you can paste the read name in by just clicking with the middle mouse button. This occurs because, when you clicked on the

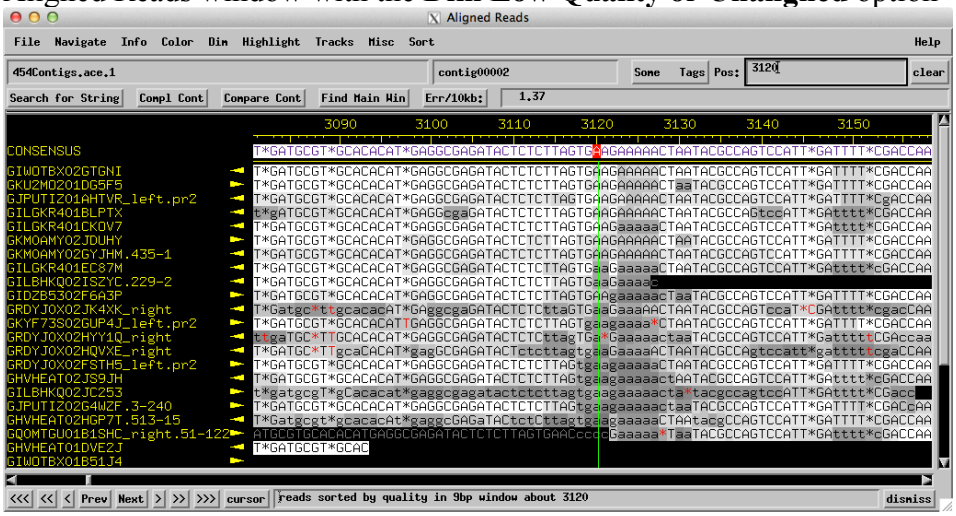
read name in the Aligned Reads Window with the left mouse button, the read name was loaded into the paste buffer.

8) DIMMING ENDS OF READS

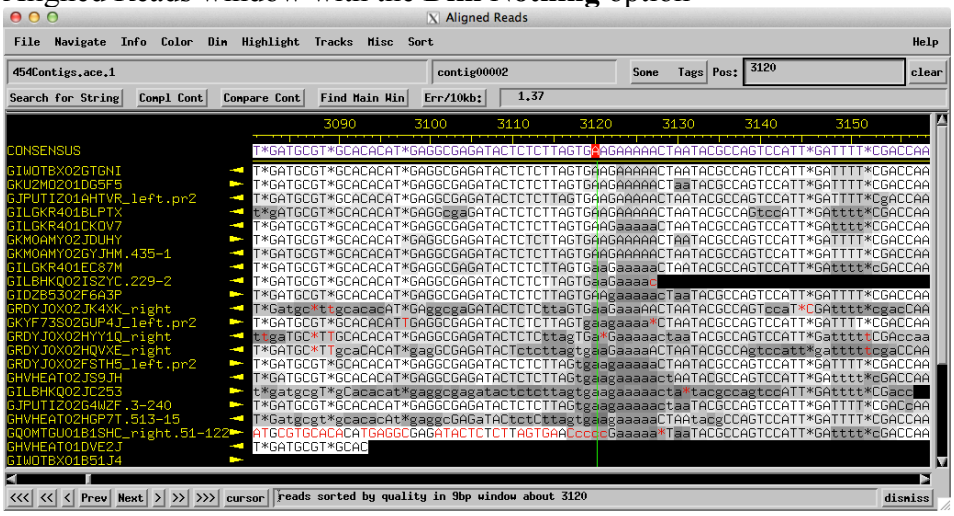
Scroll so that location '3120' is about in the middle of the Aligned Reads Window. Push the left mouse button down on the top bar menu item '**Dim.**' A list of choices will then appear. Drag the cursor down to select '**Dim Nothing**' and release.

Now look at what has happened to the color of the bases. The ends of the reads that used to be shown with a black background now appear red with a grey background. You are seeing the clipped-off bases with all the same information as any other base. Since there are large numbers of red (discrepant) bases, the screen becomes distracting and busy. Thus by default the low quality clipped-off bases are shown with a black background and a grey foreground so they don't distract you. Make this change by selecting '**Dim Low Quality or Unaligned**.'

Aligned Reads window with the **Dim Low Quality or Unaligned** option



Aligned Reads window with the **Dim Nothing** option



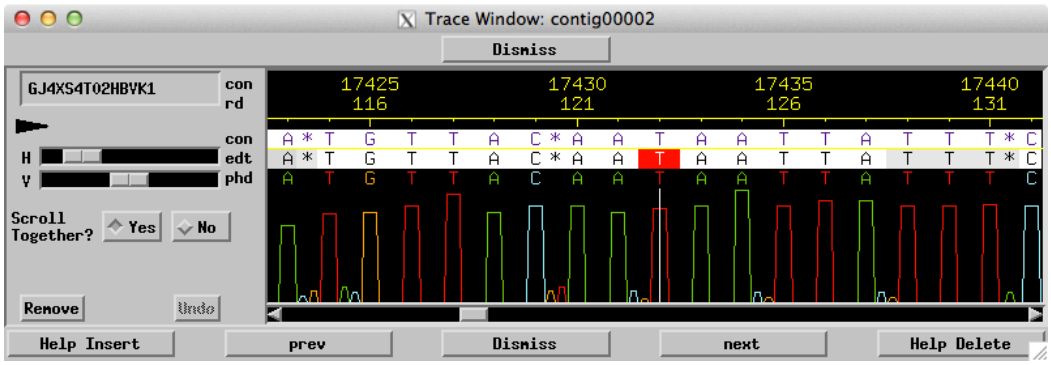
Notice there is a distinction here between '**low quality ends of reads**' and '**unaligned ends of reads**'. Unaligned ends of reads can be low quality as well, or they can be high quality. A read with unaligned high quality ends could indicate that the read has been misplaced in the assembly or that the vector or adapter sequences have not been clipped properly. In Sanger sequencing, another common cause of unaligned high quality ends of reads is chimeric reads (i.e. subclones containing two unrelated inserts).

Point with the cursor to a read name and hold down the right mouse button. You will notice there is a line that says "**high quality from nnn to nnn; aligned region from nnn to nnn; chem: 454**". This is giving the same information in number form. Highlight the read name first (see HIGHLIGHTING READ NAMES above) so you don't lose the read as you scroll. Then check to see that the numbers agree with the dimming.

You can play with the dimming options a bit. Then return it to '**Dim Low Quality**' for the rest of this tour.

9) TRACES

Point with the mouse at a base in one of the reads and click with the middle mouse button. The Trace Window showing the traces for that stretch of reads should pop-up.



There are two rows of numbers:
'**con**' are the consensus positions, the positions in the contig
'**rd**' are the read positions

There are three rows of bases in the trace window:
'**con**' is the consensus
'**edt**' is where you can edit the base calls of the read
'**phd**' is the original base calls

Notice that a red rectangle blinks (the 'cursor') in the corresponding positions of the Aligned Reads Window and the Trace Window.

10) EDITING

Try editing in the Trace Window. You can click the left mouse button on a base in the '**edt**' line to set the cursor (a blinking red rectangle). You can directly overstrike a base by typing a letter.

Try this. Try undoing it (by clicking on '**undo**' on the left hand side of the trace). If you want to undo more than one edit, you will have to go back to the Consed Main Window and click on the button labeled **'Undo Edit...**'--you will learn that later.

You can overstrike with the following characters: acgt (bases), * (a pad, in effect deleting the base), and mrwsykvhdb (IUB ambiguity codes).

You can move left and right with the arrow keys below the trace.

11) INSERTING PADS

You can insert a column of pads by pushing the **space bar**. Try this. (You may need to click on a base on the '**edt**' line first.) A pad in Consed is represented by the '*' character. A pad is used to align two or more sequences such as these:

```
gttgacagtaatcta
gttgacataatcta
```

in which one sequence has an inserted or deleted base with respect to the other. By inserting the pad character, it is possible to get a good alignment:

```
gttgacagtaatcta
gttgaca*taatcta
```

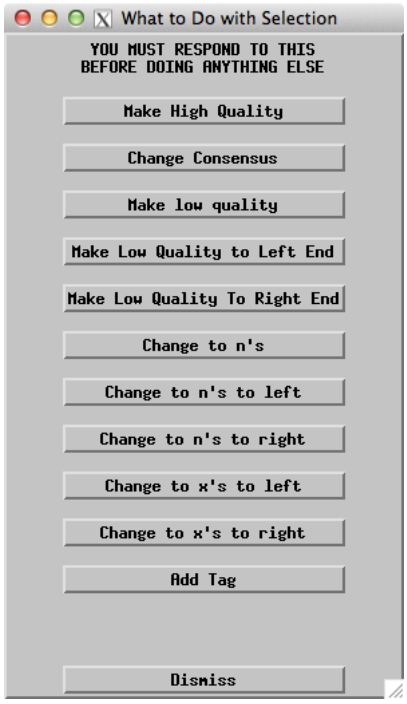
This is the purpose of pad character – it is just a placeholder. You can then overstrike a pad with a base. In this way you can insert a base, and still preserve the alignment. Note that you can only remove a pad with '**undo.**'

12) HIGHLIGHTING READS

Try highlighting a stretch of a read on the '**edt**' line by holding down the middle mouse button and dragging the cursor over some bases in the Trace Window. They will turn yellow as you drag. Then release the mouse button. A window will pop up giving you some choices of what to do with those (yellow) bases. Many of these choices will change the quality of a region and might affect the assembly when you reassemble a region with phrap.

Notice the warning at the top of the window: "YOU MUST RESPOND TO THIS BEFORE DOING ANYTHING ELSE". **This pop-up is made so that nothing else works until you choose something**. If Consed appears to be frozen, look for this pop-up window (which may

be hidden behind other windows) and dismiss it. For now, try each of these choices in the popup window, except for tags, which you'll try below.



Make High Quality- changes the quality value of the highlighted bases to 99 and adds an orange tag to the bases. This tells phrap (when it reassembles) that you are sure of the sequence here.

Change Consensus- changes the quality value of the highlighted bases to 99 and changes the consensus to agree with the selected bases.

Make Low Quality--makes the highlighted bases edited to ones of low quality. This tells phrap (when it reassembles) that you are not sure of the bases here and phrap can go ahead and make a join even if the bases in this region don't match perfectly.

Make Low Quality to Left End--same as above, but edits the bases all the way to the left end of the read.

Make Low Quality to Right End--same as above, but edits the bases all the way to the right end of the read.

Change to n's--Changes the highlighted bases to n's, which means they are unknown bases. This tells phrap (when it reassembles) to not make any join based on these bases. It is useful when you believe the bases may be in a chimeric portion of a read.

Change to n's to Left--same as above but extending to the left end.

Change to n's to Right--same as above but extending to the right end.

Change to x's to Left--Change the highlighted bases, extending to the left end, to x's, which means they are vector sequence. This tells phrap to ignore these bases for the purpose of determining overlap.

Change to x's to Right--same as above but extending to the right end of the read.

Add Tag--allows user to add any tag to a stretch of read bases.

Dismiss--you decided you don't really want to do anything with this stretch of bases.

'**Change Consensus**' has an additional function--if a read extends out on the right beyond the end of the consensus, you can extend the consensus by using this function. You might want to do this, for example, if **Crossmatch** did not correctly find the cloning site and thus clipped too much off the end, thinking it was vector. You can add these bases to the consensus by using '**Change Consensus**'. In order to extend the consensus of the contig, the read must be on the right end of the contig (this may require complementing), and the user must start the highlighting from the first base off the end of the contig.

However, if you are not going to reassemble the project, you might want to just leave the quality values the way the base caller originally called them. You can do this by using a Consed parameter (consed.extendConsensusWithHighQuality), see the CONSED CUSTOMIZATION section in the complete Consed documentation for details.

13) DELETING BASES

To delete a base, overstrike it with a '*' character. (cross_match ignores '*', so this is the same as deleting the character.) If you overstrike all bases in a column with * characters so the entire column consists of *'s (including the consensus base), there is no way to remove the column. This is OK since when you export the consensus (try the exercise on EXPORTING THE CONSENSUS below), the *'s are not exported. While you are editing in Consed, we believe there should be a visual indication that a base was deleted.

14) SAVING THE ASSEMBLY

To save the assembly, pull down the '**File**' menu on the Aligned Reads Window, and release on '**Save assembly**'. A box will pop up with a suggested name. We recommend that you always use the name it suggests. The idea is that the ace files, named

(project).454Contigs.ace.1
(project).454Contigs.ace.2
(project).454Contigs.ace.3
(project).454Contigs.ace.4
(project).454Contigs.ace.5

are in numerical order of how old they are. If you feel you are taking up too much disk space, then start deleting the ace files starting at the oldest. We do not recommend that you overwrite existing ace files. The version numbers just keep growing, and that is not a problem.

15) EXPORTING THE CONSENSUS

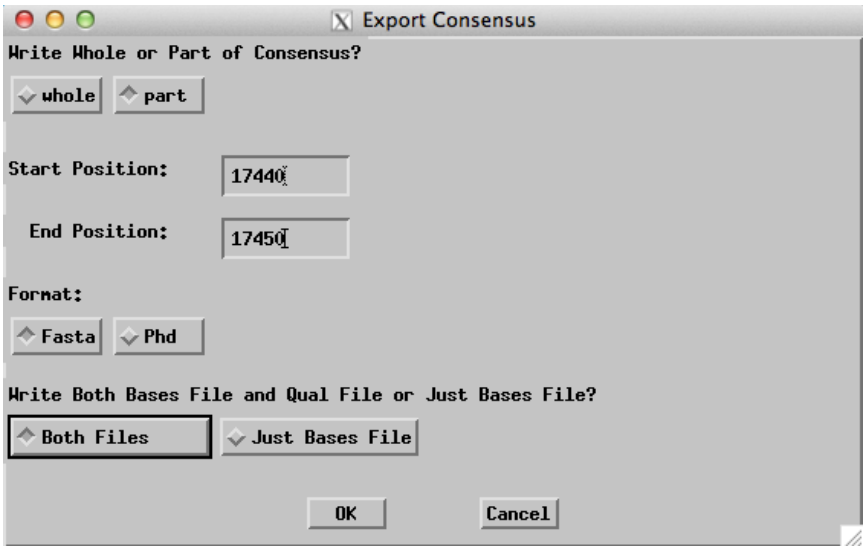
Bring the Aligned Reads Window into view again. Hold down the left mouse button on the '**File**' menu and release the button on '**Export consensus sequence**'. Notice that the consensus will be stored (in this case) in a file called 'contig00002.fasta'. Click '**OK**'. There is now a file in your edit_dir directory called 'contig00002.fasta' that contains the consensus sequence. You can examine the sequence file using any text editor (e.g. TextEdit on the Mac, vi / Emacs on Linux). You can also examine the sequence within the xterm by launching a new xterm from the Applications menu and type the following commands:

```
cd ~/Standard3/edit_dir

more contig00002.fasta
```

16) SELECTIVE EXPORTING OF CONSENSUS

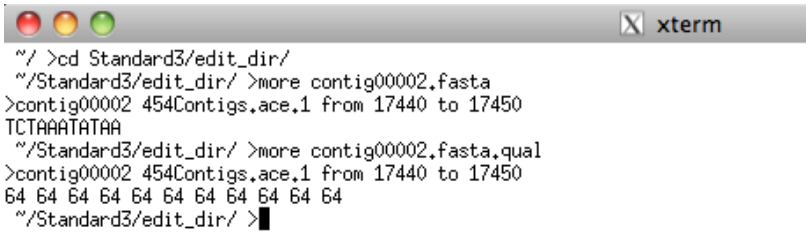
Bring the Aligned Reads Window into view again. Hold down the left mouse button on the '**File**' menu (top bar) but this time release on '**Export consensus sequence (with options)...**'. Just export a little snip of the consensus, from 17440 to 17450 bps. (You will notice this contains a pad '*' character) Under "**Write Both Bases File and Qual File or Just Bases File?**" click "**Both Files**"



Click '**OK**'. Consed will want to call this file 'contig00002.fasta' again. You can overwrite the existing file. Open these files in a text editor (e.g. TextEdit, vi, Emacs) or look in your other xterm window at these files by typing the following commands:

```
more contig00002.fasta
more contig00002.fasta.qual
```

Note that one file contains the bases (but no '*' pads) and the other contains the corresponding quality values of those bases.



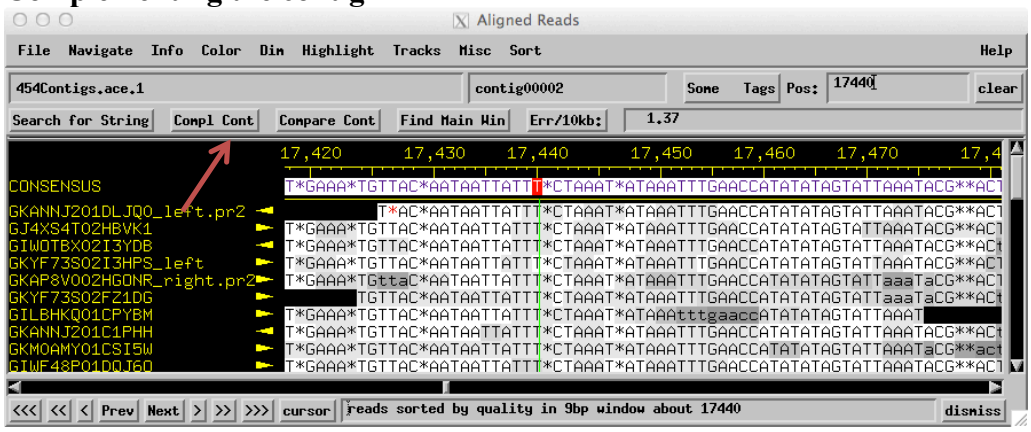
17) EXPORTING THE CONSENSUS OF ALL CONTIGS AT ONCE

Go to the Main Consed Window. Point to '**File'**, hold down the left mouse button, and release on '**Write all contigs to fasta file**'. You then can choose a filename that all contigs will be written to.

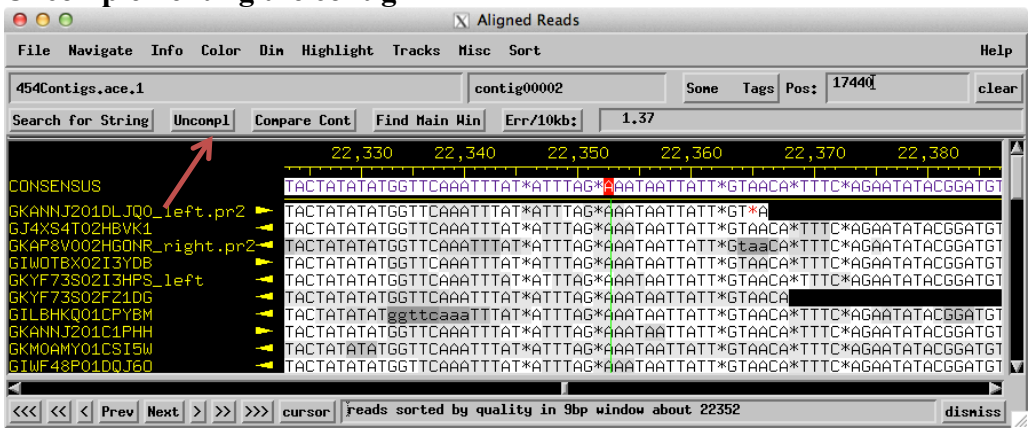
18) COMPLEMENTING THE CONTIG

Push '**Compl Cont**' in the Aligned Reads Window to complement the contig. This displays the opposite strand of the contig including the consensus and all reads**. Push this button again to uncomplement the contig**. This function can be useful when assembling contigs.

Complementing the contig



Uncomplementing the contig



19) COLOR MEANS EDITED AND TAGS

(For this step, first click on the '**Dim**' menu and release on '**Dim Nothing.**') Point to the '**Color**' menu in the Aligned Reads window, hold down the left mouse button and release on '**Color Means Edited and Tags**'.

Notice that the bases that you have edited (make sure you have edited some bases) will stand out in either white or grey (depending on whether the base was made high quality or low quality). Observe this both in the Trace Window and the Aligned Reads Window. This color mode is useful if you are interested in easily spotting which bases have been edited.

Return to the 'Color Means Quality and Tags' color mode by the following actions: point to the '**Color**' menu, hold down the left mouse button and release on '**Color Means Quality and Tags**'. Return the '**Dim**' setting to '**Dim Low Quality**.'

20) FIND MAIN WINDOW

In the Aligned Reads Window click on the '**Find Main Win**' button. This will cause the Consed Main Window to pop up in the event you have buried it under other windows or

iconified it. (This may not work with some settings of your X emulator. In that case you will have to find and click on the Main Window to bring it up.)

21) MULTIPLE UNDO EDIT

Now that the Consed Main Window is visible, click the '**Undo Edit...**' button. There will be a popup indicating the most recent edit. (If it says "no edits so far", then bring up a trace and make several edits. Then click on 'Undo Edit...' again.)

Click '**undo**'. This will allow you to see the edit that was done before the most recent edit shown above.

Click '**undo**'. You can continue undoing if you like.

You now know how to undo more than one edit. You cannot choose which edits to undo and which to not to undo – edits can only be undone in precisely the reverse order from the order in which you made them. Once you save the assembly, you cannot undo prior edits.

22) SCROLLING TRACES AND ALIGNED READS TOGETHER

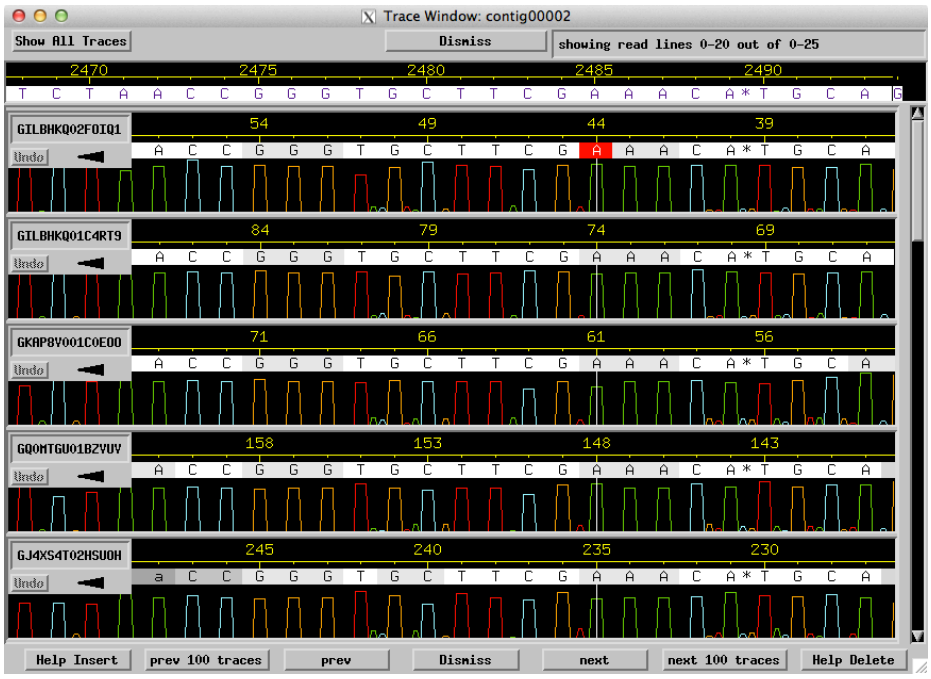
In the Aligned Reads Window, scroll along the contig to a different point. Click the left mouse button on a read whose trace is already up. Notice that the existing trace instantly scrolls to the corresponding location.

Now go to the Trace Window and scroll the traces to a new location. Click on the '**edt**' line with the left mouse button. You will notice that the Aligned Reads window will instantly scroll to the corresponding location. Thus you can keep the Aligned Reads window and the traces scrolled to the same location.

23) EXAMINING ALL TRACES

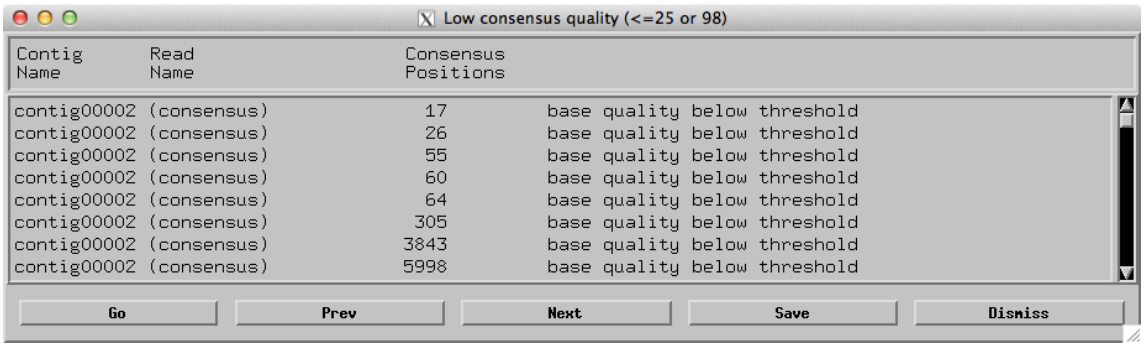
Go to a region where there are lots of reads, say base '**2480**'. Push down the right mouse button in the aligned reads window, and a pop-up menu will appear; in the pop-up menu, release on '**Display traces for all reads**'. You will see all traces displayed in a scrolling window. You can drag the scrollbar on the right down and up to see all the traces. This feature is particularly useful for polymorphism/mutation detection work.

In this Traces Window, point at one of the bases of one of the reads and click with the left mouse button. The base should start blinking in red. Now push the down arrow key on your keyboard. The cursor should move to the next read. Repeatedly press the down arrow key. The display will scroll so that you can continue to see the base the cursor is on in the different reads. Try the up arrow key as well.



24) NAVIGATING LOW CONSENSUS QUALITY

Low consensus quality is used to identify bases that have a high probability of error. In the Aligned Reads Window, pull down the '**Navigate**' menu and release on '**Low consensus quality**'. You will see a list of locations. Move the 'Low consensus quality' window down so that you can see the Aligned Reads Window.



Repeatedly click on the '**Next**' button until you reach the end of the list. (Low consensus quality means an area in which the bases each have too high a probability of being wrong.) This saves you from having to look through large amounts of high quality data trying to find problem areas.

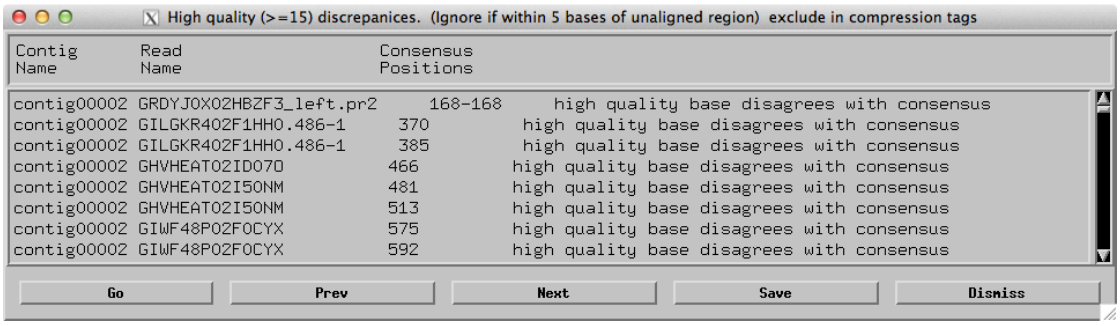
There are two 'Next' buttons – one on the Aligned Reads Window and one on the Low Consensus Quality Window. You can click on either, but it is probably more convenient to use the 'Next' button on the Aligned Reads Window. Thus you can keep the Aligned Reads Window in front with input focus and keep the Low Consensus Quality Window pushed out of the way.

In our experience, this will be one of the most important navigation lists you will use. In fact, a major part of sequence improvement is to add reads to these low quality areas until this list is reduced to nothing.

25) NAVIGATING HIGH QUALITY DISCREPANCIES

'**Dismiss**' the Low Consensus Quality Window. Pull down the '**Navigate**' menu again and release on '**High quality discrepancies as above, but omitting tagged compressions and G_dropouts**'. By default, Consed treats discrepancies with a quality score of 40 or above as high quality discrepancies. However, we can modify the score threshold for high quality discrepancies through the **General Preferences** window.

First, '**Dismiss**' the High Quality Discrepancies Window. Click on '**Find Main Win**'. In the Consed Main Window, pull-down the '**Options**' menu and release on '**General Preferences**'. Notice that the default for 'Threshold for High Quality Discrepancy (lowest high)' is '40.' Change it to '**15**' and click '**Apply & Dismiss**'.

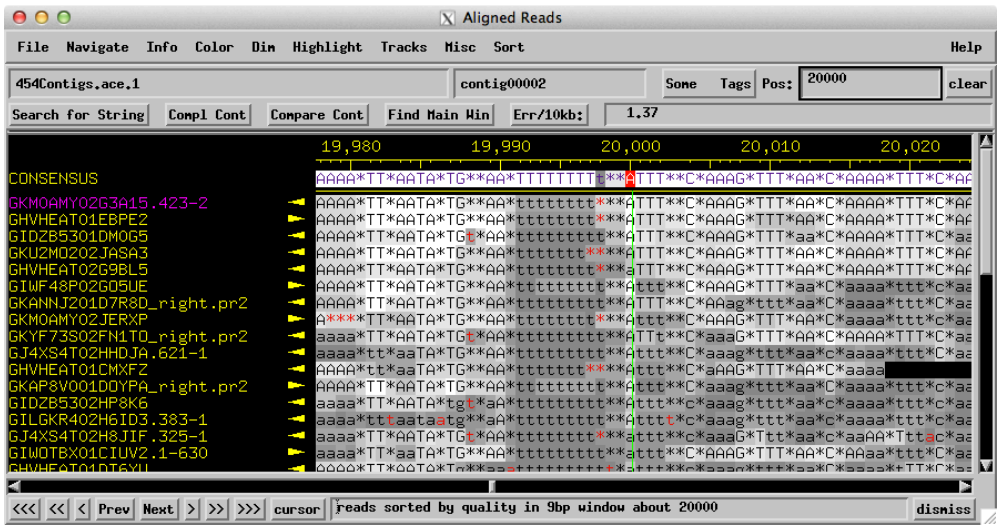


Then follow the steps above to bring up the High Quality Discrepancies Window. Now you will notice that Consed labels many more discrepant bases as high quality discrepancies. Click '**Next**' repeatedly to go successively to the next high quality discrepancy in the Aligned Reads Window.

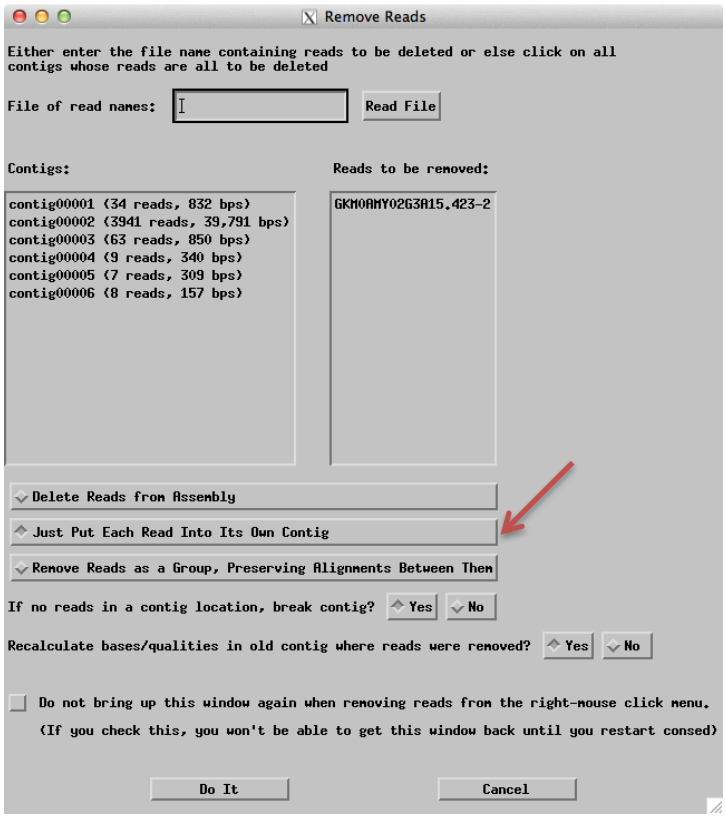
You can also double click on a particular line in the High Quality Discrepancies Window to go to that location. Alternatively, you can single click on a line and then click the '**Go**' button. '**Dismiss**' the High Quality Discrepancies Window. Reset the '**Threshold for High Quality Discrepancy**' to 40.

26) REMOVING READS

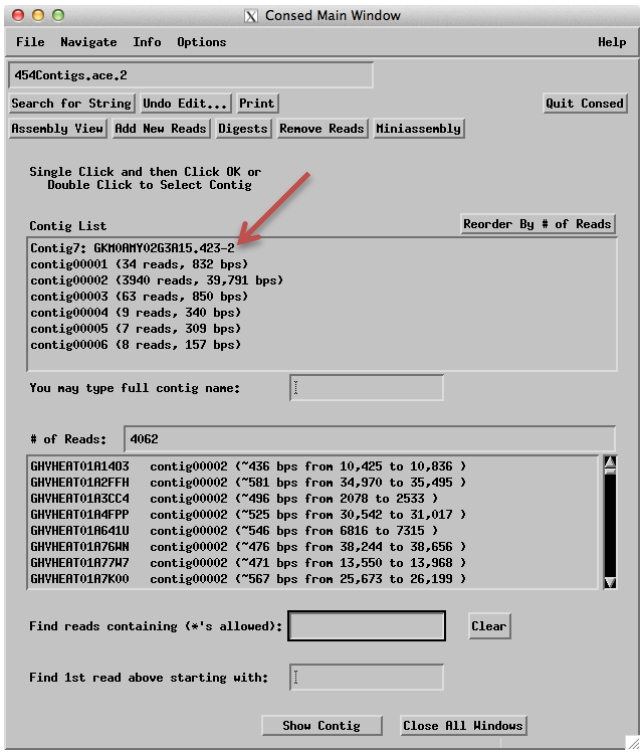
It is possible that the computer has assembled one or more reads in the incorrect place. This is often indicated by the presence of many high quality discrepancies. Navigate to position **20000** in the Aligned Reads. Point to the read name of read **GKM0AMY02G3A15.423-2** and hold down the right mouse button. Release on '**Remove read GKM0AMY02G3A15.423-2 from this contig**.' (If you cannot find this read then you may have missed the instruction of uncomplementing the contig in step 18 above.)



The “Remove Reads” dialog will appear and we can see the read we have previously selected under the “**Reads to be removed**” section. Select the option “**Just Put Each Read Into Its Own Contig**” and then click on the “**Do It**” button.



A “Reads Removed” dialog box will appear indicating that the read has been removed from the contig. The read is put into its own contig (Contig7) and the old contig (2) is redrawn without the read in it. At this point you should save the assembly – **you should always save the assembly after removing reads**. (See item 14.)



27) ADDITIONAL NAVIGATORS

Similarly, try the other navigate lists:
Edits
Unaligned high quality regions (this list will be empty with this data set)
Regions covered by only 1 strand and only 1 chemistry
Regions covered by only 1 subclone

Unaligned high quality regions are regions in which the traces are high quality so there is no question of the base calling, but the region differs so much from other reads that phrap has given up trying to align the region with the consensus. This could be due to a chimeric read, or perhaps the read belongs somewhere else in the assembly.

We believe that regions should be **covered by at least two subclones** to prevent the possibility of there being a deletion in a single subclone.

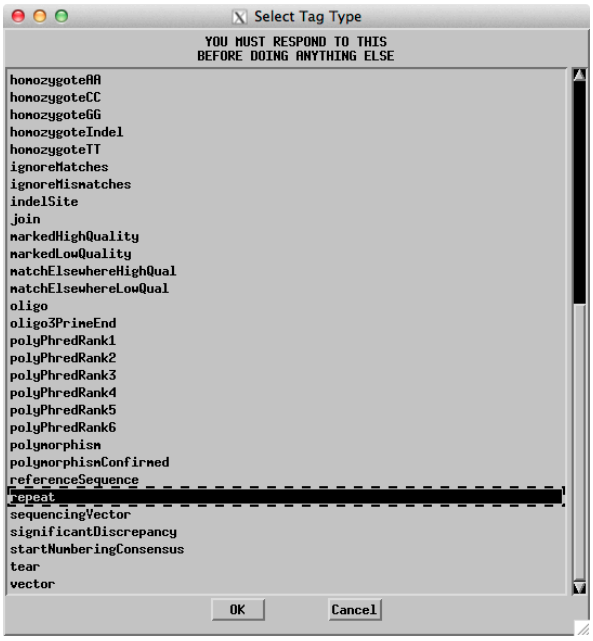
There are so many different problem lists that you may forget to check one of them and thus miss a serious problem. Thus we combined them all into a single list. This is the first menu item: '**LowCons/HighQualDiscrep/SingleStranded/SingleSubclone/UnalignedHigh**'. This list is most useful when working on projects based on Sanger sequencing, for working on 454 assemblies other navigators found on the Consed Main Window will be used instead.

28) NAVIGATING BY TAGS

Also try to navigate by tags by selecting '**Tags**' under the '**Navigate**' menu: when the Select Tag Type Window appears, double click on '**repeat**'. **(Note that you must respond to this**

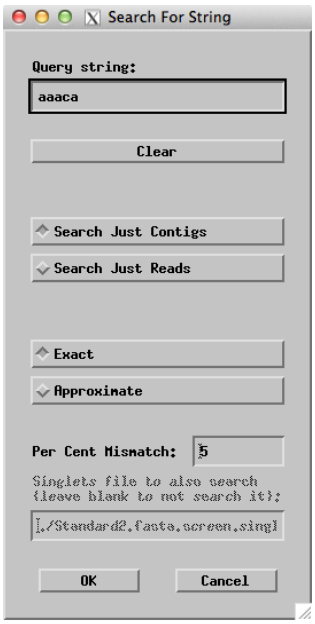
window before you can interact with other Consed windows) A new navigation list will appear that shows the positions of all the repeat tags in the contig (i.e. contig00002).

There is also a way of getting a list of a particular tag type in all contigs. Click on '**Find Main Win**'. In the Consed Main Window, point to the '**Navigate**' menu, hold down the left mouse button, and release on '**Tags in all contigs**'. Continue as in the previous step.



29) SEARCH FOR STRING

Try the '**Search for String**' button (left side of the Aligned Reads Window). Type in a string of bases (such as '**aaaca**') and click '**OK**'.



There should be a list of locations where the sequence **aaaca** are found. Double click on one of the hits (or single click on it and click on '**Go**'.) Notice that the Aligned Reads Window scrolls to that position and has the cursor on the found string. (It might be complemented.)

'**Dismiss**' this window. Try this again, only this time in the Search For String Window select '**Search Just Reads**'. Then click '**OK**'. You will notice there are many more hits. This is because this option shows hits in each read, even if they are at the same consensus position.

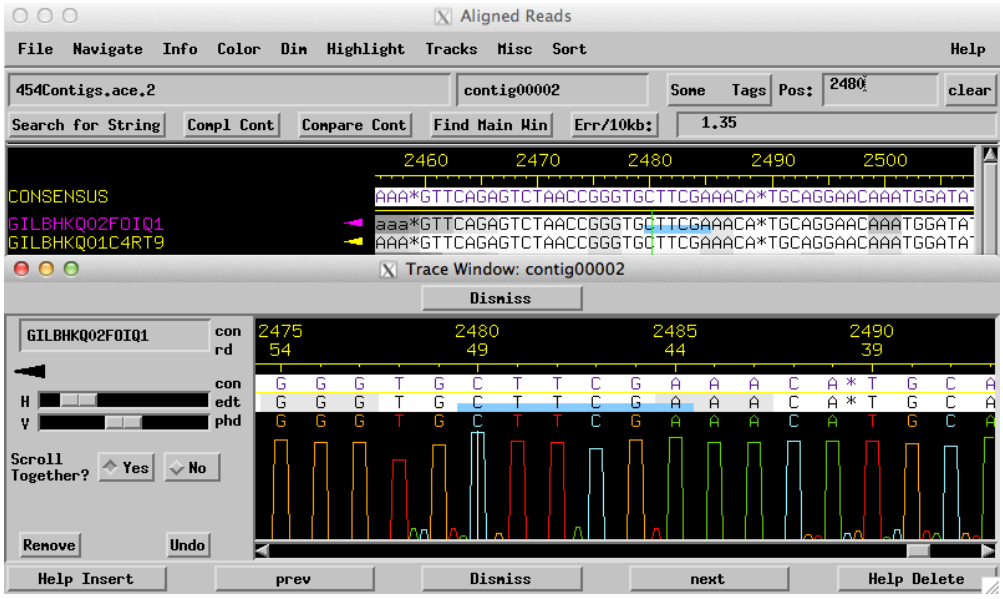
You can also search for approximate matches by clicking on '**Approximate**' instead of 'Exact' in the 'Search For String' dialog. Note that the 'Per Cent Mismatch' box only applies to approximate match searches.

30) COPY AND PASTE

In the Aligned Reads Window, swipe some bases by holding down the left mouse button as you mouse-over a string of bases. You should see the bases turn yellow, at least temporarily. Then click the '**Search for String**' button. Use the middle mouse button to paste the bases you have just swiped into the 'Query string:' box. Notice that you can swipe bases either from the consensus or from a read. The Search for String program is case-insensitive so the pasted bases can be either in upper or lower case.

31) CREATING TAGS

Bring up a trace for a read and swipe some bases on the '**edt**' line while holding the middle mouse button down. A list of choices will pop up. Select '**Add Tag**'. Type in a comment in the box at the bottom, and select '**comment**' from the list of tag types. Click on 'OK.' A light blue box will appear on that read in both the Aligned Reads Window and the Trace Window.



To see the comment, just point to the blue box in the Aligned Reads Window and you will see the comment in the lower right hand corner of the Aligned Reads Window. Alternatively, you can click on that blue tag in the Aligned Reads Window with the right mouse button and release on the '**Tag: comment Show more info**?' item. Alternatively, you can click on the blue tag in the Traces Window with the right mouse button.

Try creating some other kinds of tags by again swiping some bases in the **edt** line of the Trace Window and then selecting a different tag type. You will notice that different tags are marked in different colors. You can always use the methods above to see what kind of tag it is if you forget what a particular color means.

Try creating some tags that overlap each other. You will notice that the overlapping region will be purple. If you want to know which tags overlap, you can use any of the methods already discussed above.

You can create tags on the consensus in the same way. In the Aligned Reads Window, use the middle mouse button to swipe some bases on the consensus in the Aligned Reads Window. A dialog box with a list of tag types will appear. Click on one of them.

Try to create other tags again somewhere else.

32) ADD NEW READS

Projects will often have problem areas that need additional sequencing data. Before we can incorporate new reads into the assembly, we must first save all of our changes. Go back to the Consed Main Window and select 'File', then select 'Save assembly' to save all our changes in a new ace file. For this project the files have already been placed in the folder. In the Consed Main Window, click on the **Add New Reads** button. A new dialog will appear

which shows the list of files ending with .fof in the "edit_dir" of your current project directory. The .fof file contains a list of filenames that corresponds to the chromatograms files we want to add. Double click on the file **'reads_to_add.fof**.'

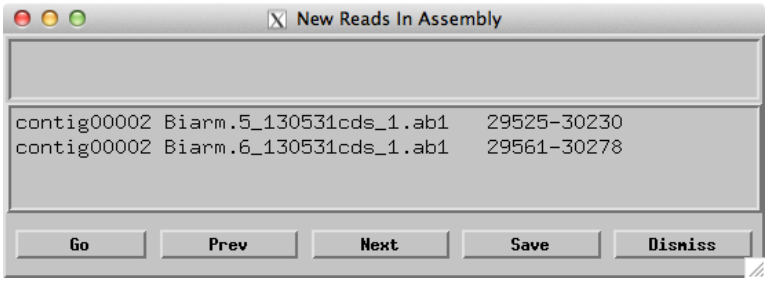
Consed will ask us two questions:

Do you want to recalculate the consensus quality values where each of the new reads is aligned?"

If a read doesn't align against any existing contig, do you want to have it go into a contig by itself?

You should answer "**yes**" to both questions.

There should be lots of progress output in the xterm window from which you started Consed. When it completes, there will be a **Reads Added** Window popup with a report of which reads were added. In this case, it should say that two reads were successfully added and list them. Save the new assembly.



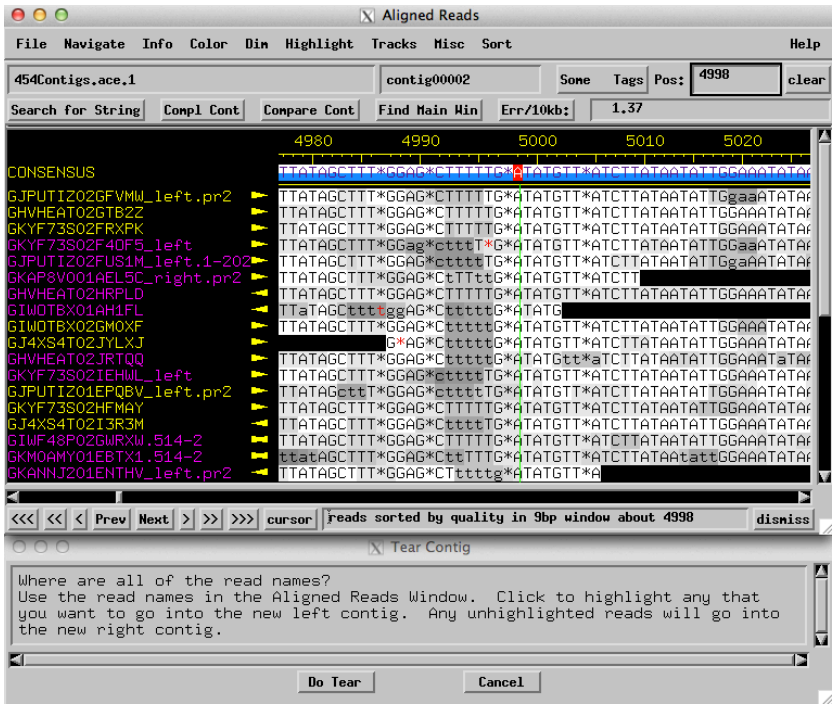
To see the new reads incorporated into the assembly select a read and click the **Go** button or double-click on the read name, either of these will take you to the beginning of the read.

33) TEAR CONTIG

Just so you get the same results as your instructor, exit Consed from the Main Window (see item 35) and bring it up again using the original ace file: "**454Contigs.ace.1**". (See item 2.)

If it asks if you want to apply edits, just say '**no**'. When there are multiple misassemblies in your project, you may want to just tear the contig apart in several places and then join the pieces back together in a different way. Let's try it! Go to location **4998** in the Aligned Reads window. Point the mouse at the consensus base at 4998 and push the right mouse button down. Release the button on '**Tear Contig at This Consensus Position**'.

Consed will highlight the reads that will go into the left contig and unhighlight reads that will go into the right contig. Leave everything as it is and just click '**Do Tear**'.

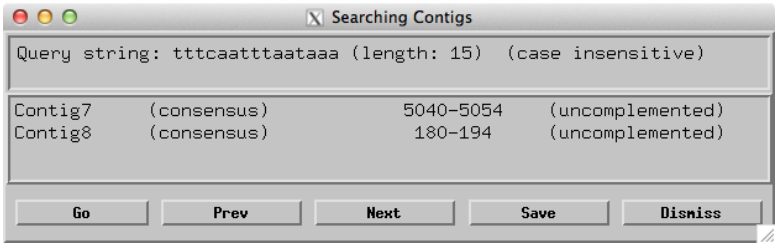


Now you should have two Aligned Reads Windows on top of each other. One should contain 'Contig7' and the other 'Contig8' (Your contig numbers may not match depending on how many of the steps above you actually completed). **Dismiss** the little window that says 'Tear Complete'. **Save the new assembly**.

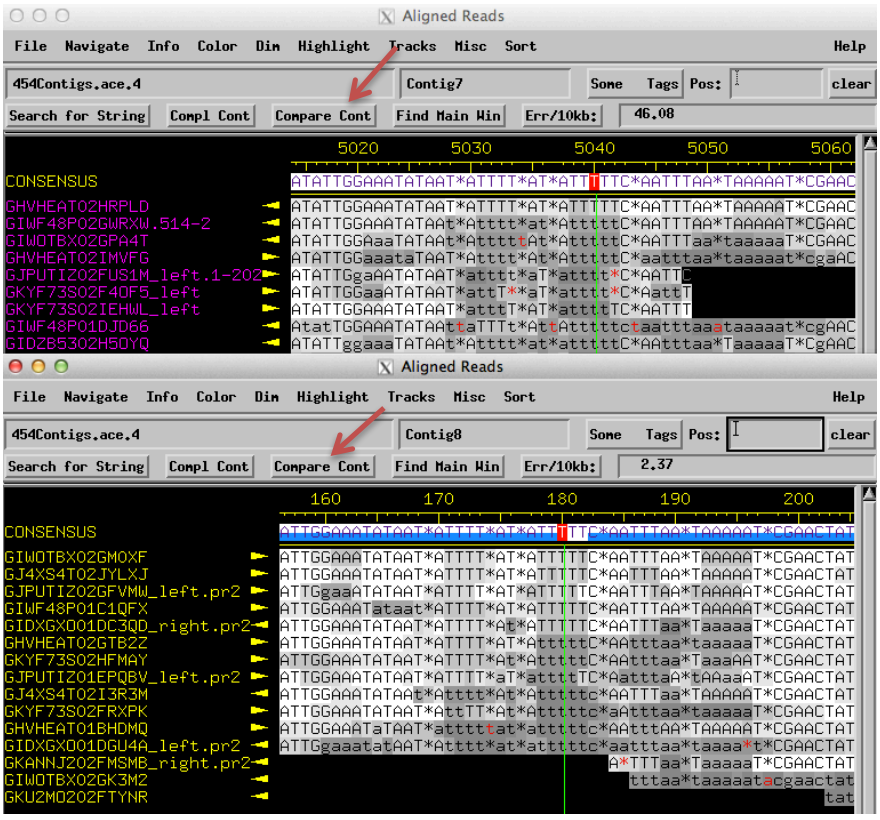
34) JOIN CONTIGS

Now let's join these two contigs back together: Click on '**Search for String**' and type in the following bases: **tttcaatttaataaa**. Click '**OK**'.

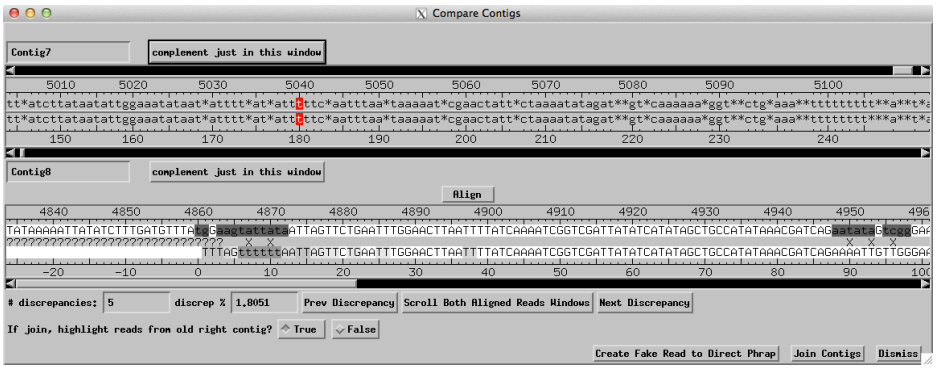
Search for string should find two locations, one in Contig7 and one in Contig8:



Double click on the first one. The Aligned Reads Window for Contig7 will scroll to location 5040 and the window will rise up. In that Aligned Reads Window, click on **'Compare Cont'**. The 'Compare Contigs' window will open.



Now go back to the Search for String results window and double click on the contig8 entry. The Aligned Reads Window for Contig8 will scroll to location 180 and it will become the topmost window. In that Aligned Reads Window, click on **'Compare Cont'**.



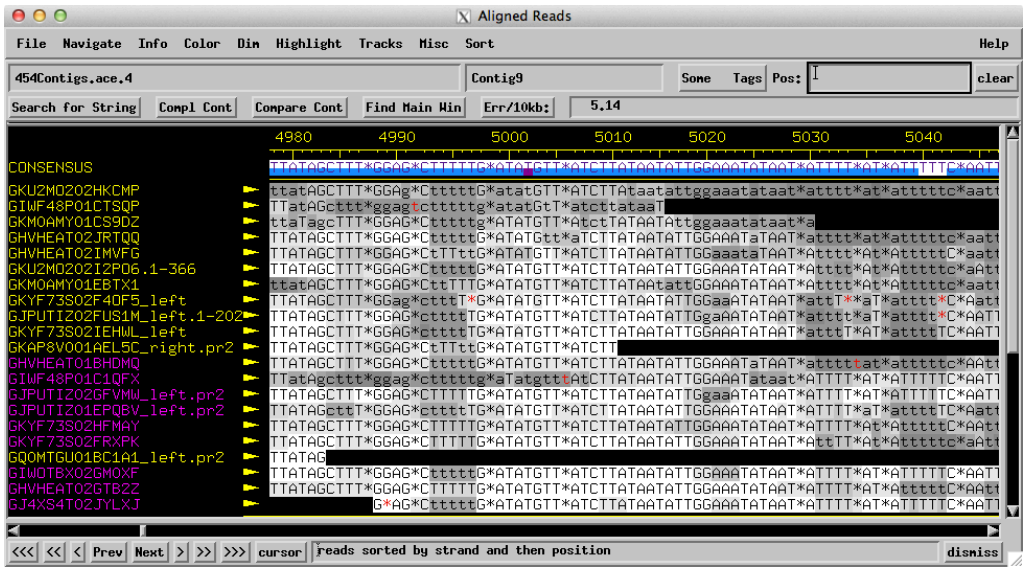
Now the Compare Contigs Window should be visible. In the Compare Contigs Window, try scrolling back and forth. You can change the cursors (blinking red), but if you do, please return them to the locations 5040 and 180 for the next step. The cursors 'pin' these bases together when doing an alignment. (The algorithm is a pinned and banded Smith-Waterman alignment.)

Click on the **Align** button. Try scrolling the alignment by dragging the scrollbar in the lower half of the Compare Contigs. An 'X' means there is a discrepancy between the two contigs. There is also a 'P' (see if you can find it!) The P indicates the bases that you pinned together.

You will also notice that some bases are lighter and some are darker. This indicates quality just as in the Aligned Reads Window. You will notice that wherever there is a discrepancy (an 'X') one of the bases is low quality. This is your cue that the discrepancy is just a base calling error rather than indicating that the two contigs really are different, but have similar (repetitious) sequences.

Click with the left mouse button on either contig in the bottom alignment. You will notice that both contigs will have the red blinking cursor in the same position.

When you scroll to the beginning of the alignment, you will notice the region is filled with question marks. The question marks indicate regions where Consed can no longer align the two contigs. If this were a real project we would carefully inspect the alignment to assess if this is a good join (all high quality bases agree) or an inappropriate join (too many high quality bases disagree). We will discuss this more in detail later. In this case since we tore this apart above and simply wish to join these two contigs we will do so by clicking the **'Join Contigs'** button. The two previous Aligned Reads Windows will disappear and there will be a new one: 'Contig9'. You have made a join! Again anytime a tear or join has been made it is important to **Save the assembly,** do that now.



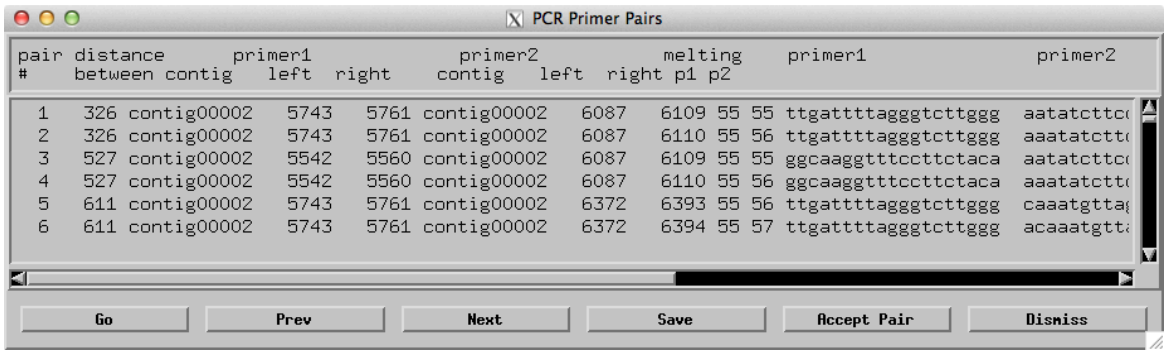
Scroll left and right in the Aligned Reads window. You will notice that many of the reads are highlighted. These are the reads that came from the previous "right" contig. To unhighlight all of these reads at once, point to the "**Highlight**" menu (top bar), hold down the left mouse button and release on "**Unhighlight All Reads in All Contigs**".

It is possible to have more than one 'Compare Contigs' windows up at a time. This allows you to investigate a repeat that has more than two copies.

35) DESIGNING PCR PRIMERS

Exit Consed from the Main Window (see item 35) and bring it up again using the original ace file: "**454Contigs.ace.1**". (See item 2.) In the aligned reads window for contig00002 navigate to position 5995. This region has a mononucleotide run that has about half of the reads with 7 T's and half with 8 T's. To resolve this region unambiguously a finisher would need additional data, preferable from a sequencing technology that does not have issues with mononucleotide runs. Consed has the ability to look for oligonucleotide primers that can be used in Sanger sequencing directly on template clones (if available) or used in PCR reactions to generate the proper template.

To design PCR primers, go to the left end of the problem area you wish to amplify, point to the consensus (base 5990), hold down the right mouse button and release on '(Top Strand) First PCR Primer'. Click "Dismiss" to remove the info window on the number of possible primers found by Consed. Then scroll to the location on the right end of the region where you want to pick the second PCR primer (base 6000). Point to the consensus, hold down the right mouse button and release on "(Bottom Strand) Second PCR Primer". There will be a pause and then a list of PCR primer pairs will appear in a new window. Click on the first (top) pair and click "Accept Pair". You will now see PCR primers (in yellow with a red tip) at 5743 and 6109.



You can modify the parameters for choosing PCR primer pairs by going to the Consed Main Window, pointing to "Options", holding down the left mouse button, and releasing on "Primer Picking Preferences." [PCR primers are screened for: melting temperature and length, the melting temperature of the 2 primers must be sufficiently close to each other, each primers must not stick to itself or to the other primer, no mononucleotide repeats, only ACGT's (no n's or ambiguity codes), and primer pair must not amplify any other location.]

36) EXITING CONSED

On the Aligned Reads Window, point to '**File**' menu, hold down the left button and release on '**Quit Consed**'. If it asks you some questions, answer '**Quit Without Saving and Discard .wrk File**'.

37) THE .WRK FILE

Consed keeps a log of all changes you make to an assembly: adding new reads, putting reads into their own contigs, making joins and tears, adding and removing tags, and changing bases. This log is kept in a file ending with ".wrk". If Consed (or your computer) crashes, Consed can use the .wrk file to recover your edits (base changes, adding and removing tags, and complementing contigs).