11/20/12 Mini lecture on dideoxysequencing

Recap of of where we are with respect to the analysis of rifR sequencing exercise

What have we done so far?

Sanger Dideoxy [Cycle] Sequencing

What does the dideoxy refer to?

How does a sequencing reaction differ from PCR ?

Sanger dideoxysequencing
http://www.dnalc.org/resources/animations/sangerseq.html

dNTP stands for:

ddNTP stands for:

dNTP = 2’ deoxynucleoside triphosphate

ddNTP = 2’, 3’ deoxynucleoside triphosphate

How does a sequencing reaction differ from PCR ?

- One primer rather than 2
- arithmetic rather than exponential copying because product is single stranded copy of one of the template strands
- original template is not being amplified
- Big dye terminators: fluorescent label and 2' 3' dideoxy

HERES A RECIPE for a DIDEOXY CYCLE SEQUENCING reaction:

ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit
(With AmpliTaq DNA Polymerase, FS)

Preparing Sequencing Reactions

1. Thaw Terminator Ready Reaction Mix on ice.

2. For each reaction, add the following reagents to a separate thin-wall PCR tube:

Terminator Ready Reaction Mix	4.0 µl
Template DNA	varies
Primer 3.2 pmol	
dH_20	varies
TOTAL VOLUME	10 µl

3. Mix well and spin briefly.

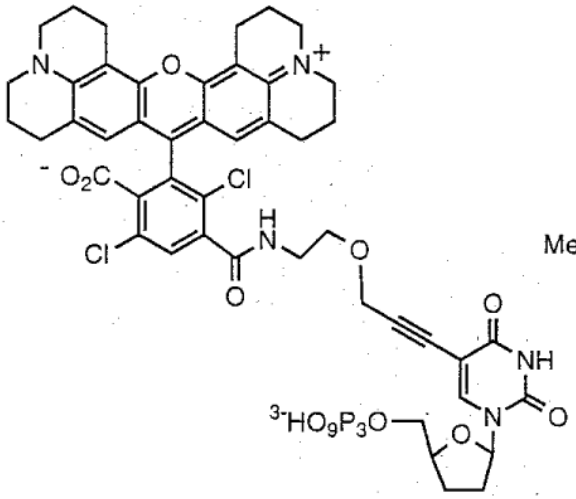
Cycle Sequencing on the Gene Amp 9600/9700

1. Place the tubes in a thermal cycler and set the volume to 10 µl.
2. Program: 25 cycles: 96°C – 10 sec (100% ramping) {The ramp speed should be 1°C /sec.}
   50°C – 5 sec (60% ramping)
   60°C – 4 min (100% ramping)
   4°C hold

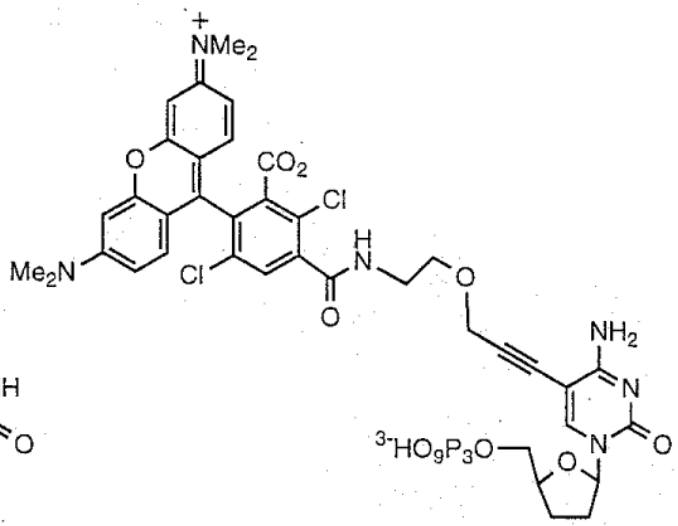
Terminator Ready Reaction Mix:

Terminator Ready Reaction Mix:

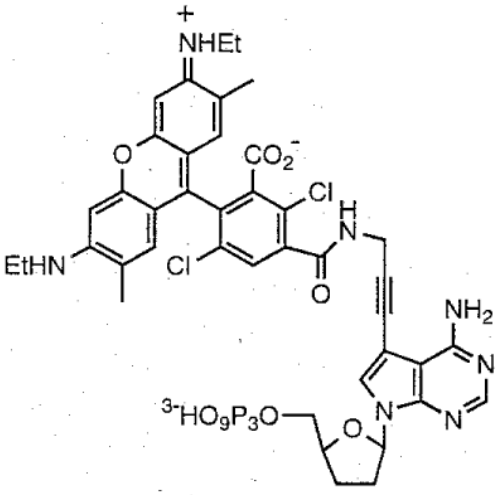
- A-BigDye Terminator v3.0
- C-BigDye Terminator v3.0
- G-BigDye Terminator v3.0
- T-BigDye Terminator v3.0
- Deoxynucleoside triphosphates (dATP, dCTP, dITP, dUTP)
- AmpliTaq DNA Polymerase, FS
- $MgCl_2$
- Tris-HCl buffer, pH 9.0



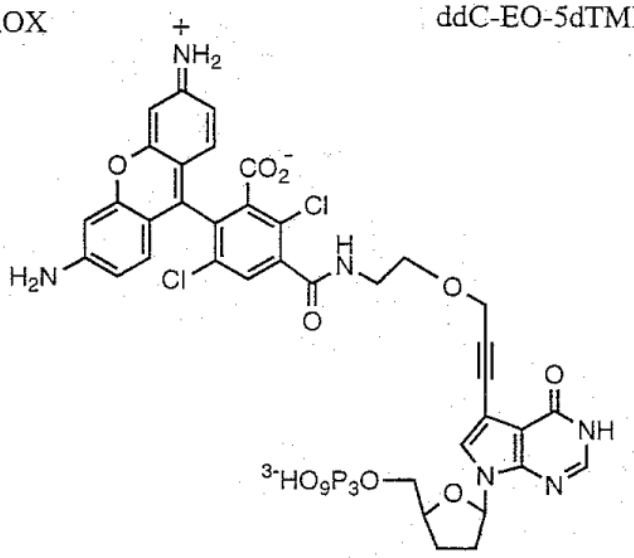
ddT-EO-6dROX



ddC-EO-5dTMR

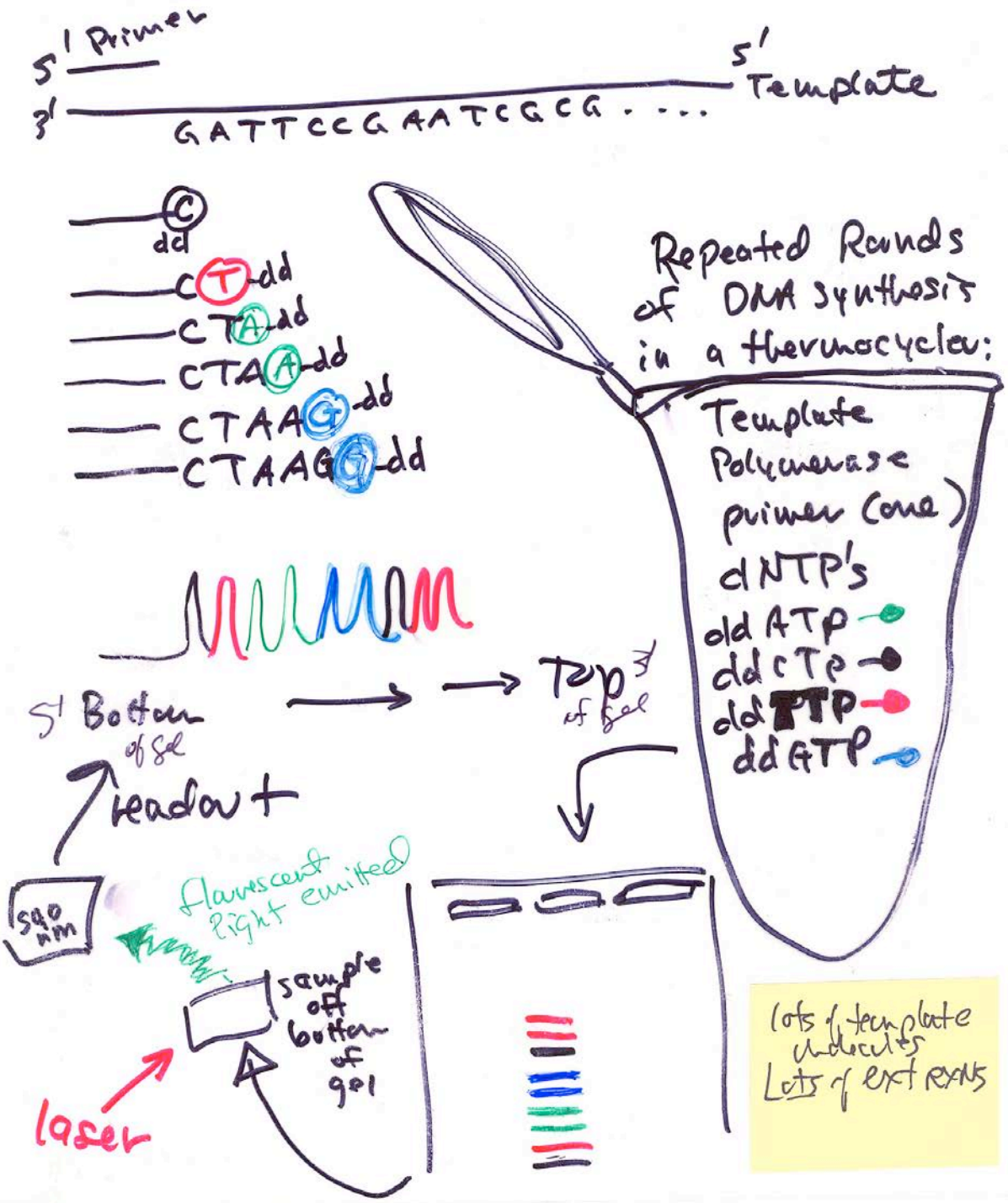


ddA-PA-5dR6G



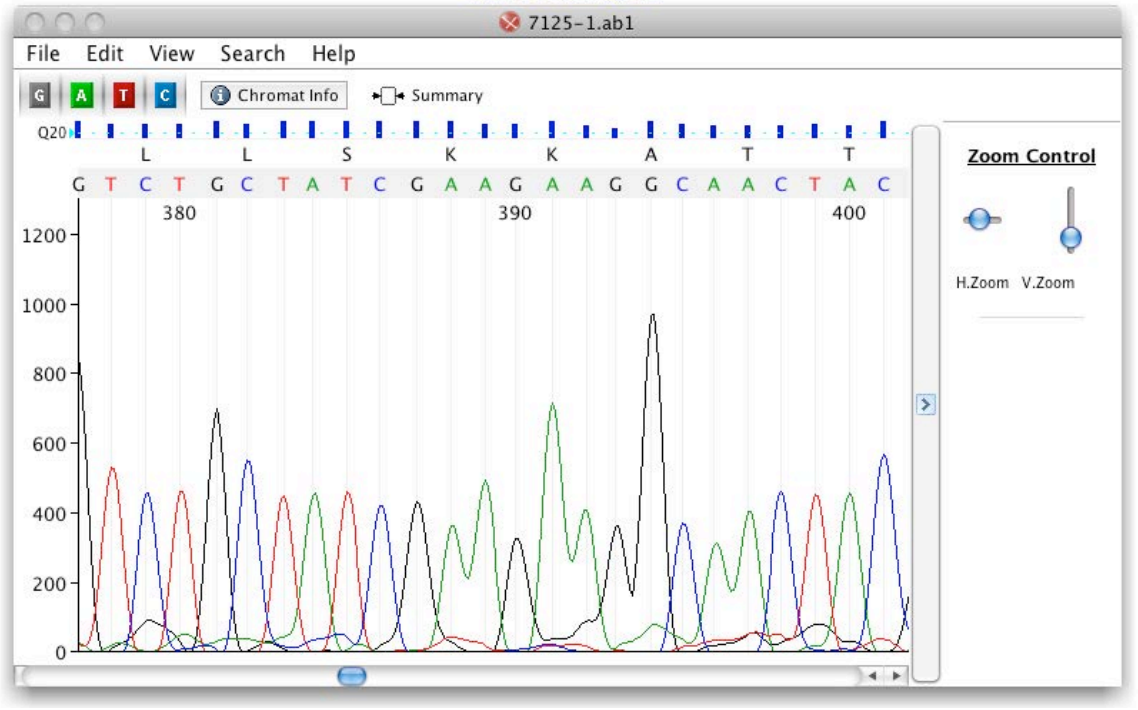
ddG-EO-5dR110

Cycle Sequencing



Display Results for 7125-1 1-rpoB-S

3730 Chromatogram



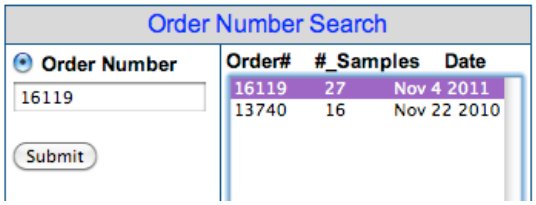
orient 5 → 3’

Go to the Nevada Genomics Web Site
http://www.ag.unr.edu/Genomics/

On the left side of page click on DNA Tools
Then click on Login to dnaLIMS
Both the login name and password are: Biol322
Click on

Download DNA Results

Then request order # 188677 or 188678



Click on View (not Download)

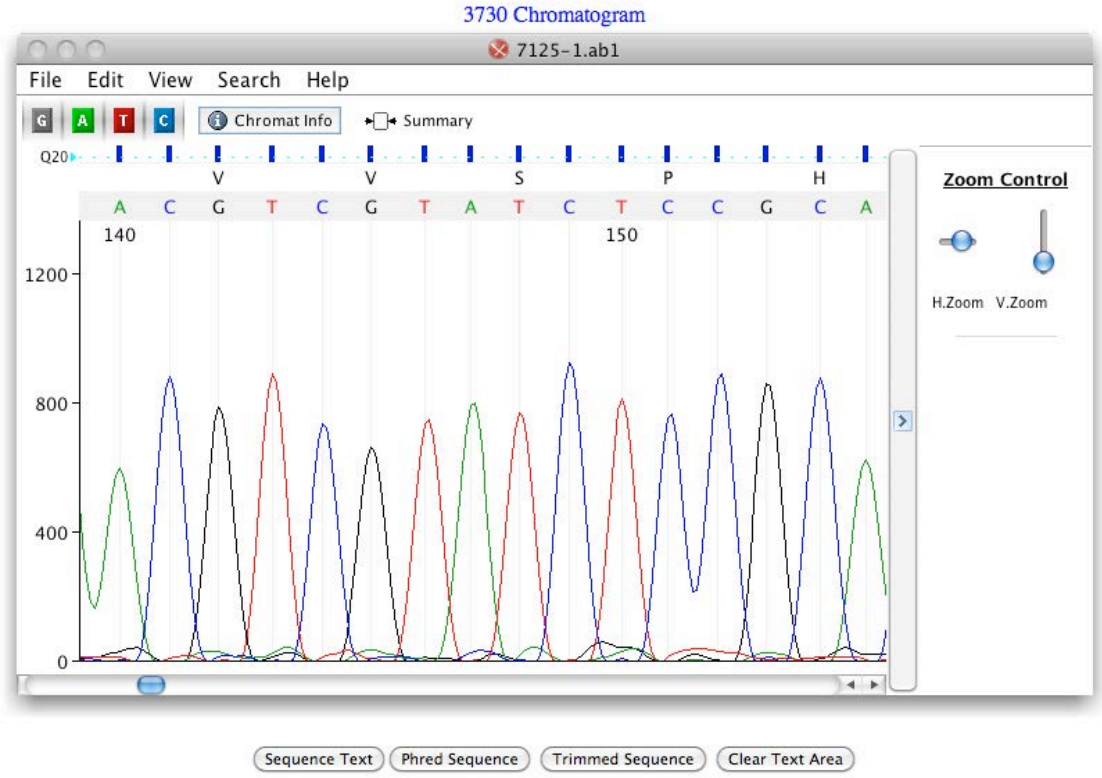
Mac: Hold down Control Key, Click Mouse. Select Save Linked File As...
PC: Right Mouse Click, Save Link As...

Order Number: 18677

Rows in grey are checked in to be rerun.

Order#	Req#	SeqId	Download	View	Sample	Primer	UID	Date	Phred Q20	Trimmed	Edited Download	Comments
18677	620582	8613-33	Text Chromat	View	1	rpoB-S	2367	Nov 6 2012	phd qual 568 fasta scf	fasta		Results Available
18677	620583	8613-34	Text	View	2	rpoB-S	2367	Nov 6 2012	phd qual 545	fasta		Results Available

Play around with H.Zoom and V.Zoom until your chromatogram looks like this:



Phred Quality Score
http://en.wikipedia.org/wiki/Phred_quality_score

Formula for a Phred score:
$Q = -10 \log_{10} P(error)$

For example, if Phred assigns a quality score of 30 to a base, the chances that this base is called incorrectly are 1 in 1000. The most commonly used method is to count the bases with a quality score of 20 and above. The high accuracy of Phred quality scores make them an ideal tool to assess the quality of sequences.

Phred quality scores are logarithmically linked to error probabilities

Phred Quality Score	Probability of incorrect base call	Base call accuracy
10	1 in 10	90 %
20	1 in 100	99 %
30	1 in 1000	99.9 %
40	1 in 10000	99.99 %
50	1 in 100000	99.999 %

Return to previous screen and under Phred Q20, click on phd and qual – how do they relate to each other?

What are Q20 values? The base calling program used in conjunction with the ABI Prism 3730 DNA analyzer is Phred, a program developed by Dr. Phil Green and Dr. Brent Ewing. Phred reads DNA sequencer trace data, calls bases, assigns quality values to the bases, and writes the base calls and quality values to output files. The quality scores are logarithmically linked to error probabilities, as shown in the table above. It has been shown that Phred's error probabilities are very accurate. The most commonly used method is to count the bases with a quality score of 20 and above, thus the "Q20" value, which indicates an accuracy of 99% for the base called.

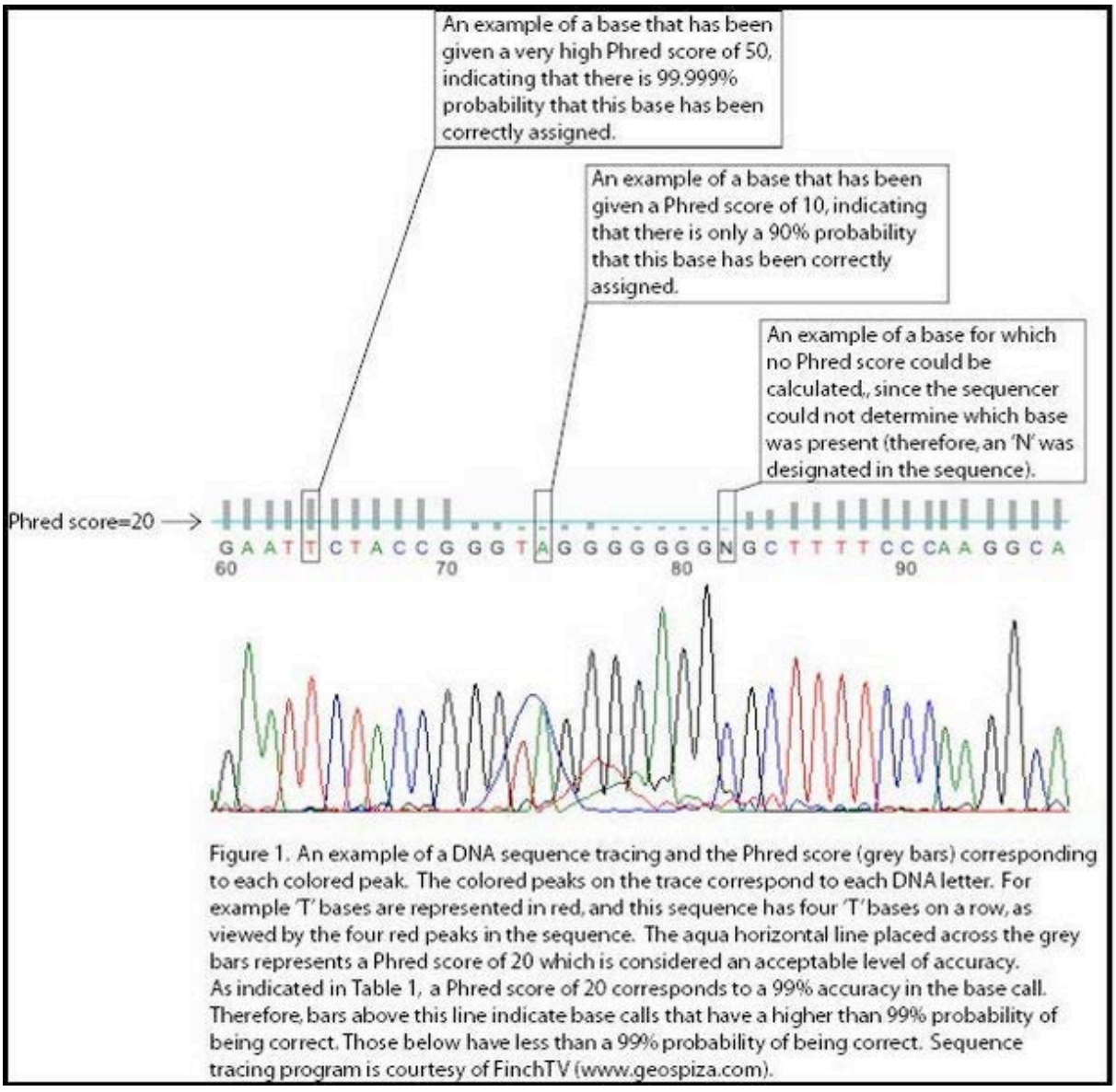


Figure 1. An example of a DNA sequence tracing and the Phred score (grey bars) corresponding to each colored peak. The colored peaks on the trace correspond to each DNA letter. For example 'T' bases are represented in red, and this sequence has four 'T' bases on a row, as viewed by the four red peaks in the sequence. The aqua horizontal line placed across the grey bars represents a Phred score of 20 which is considered an acceptable level of accuracy. As indicated in Table 1, a Phred score of 20 corresponds to a 99% accuracy in the base call. Therefore, bars above this line indicate base calls that have a higher than 99% probability of being correct. Those below have less than a 99% probability of being correct. Sequence tracing program is courtesy of FinchTV (www.geospiza.com).

click on fasta under Phred Q20

The Fasta format is as below

- name of file is preceded by a >
- sequence starts on the next line
- note that rpoB-S is the name of the primer used for the sequencing reaction
-

```
>1-rpoB-S 843 47 644 0.05
CTACCCAGCCCGGGGTAGCTGATGCCTCAGGAAATGATCAACGCCAGCCGATTTCCGCAGCAGTGAAAGAGTTCTTCGGTTCCAGCCAGCTGTCTCA
GTTTATGAACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTC
GAAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGT
ACGCACAGACTAACGAATACGGCTTCCTTGAGACTCCGTATCGTAAAGTGACCGACGGTGTTGTAACTGACGAAATTCACTACCTGTCTGCTATCGA
AGAAGGCAACTACGTTATCGCCCAGGCGAACTCCAACCTGGATGAAGAAGGCCACTTCGTAGAAGACCTGGTAACTTGCCGTAGCAAAGGCGAATCC
AGCTTGTTCAGCCGTGACCAGGTTGACTACATGGACGTATCCACCCAGCAGGTGGTATCCGTCGGTGCGTCCCTGATCCCGTTCCTGGAACACGATG
ACGCCACCGTGCATTGATGGGTGCGAACATGCAACGTCAGGCCGTTCCGACTCTGCGTGCTGATAAGCCGCTGGTTGGTACTGGTATGGAACGTGCT
GTTGCCGTTGACTCTGTGTAACTGCGGTAGCTAACGTGCTGTGTCGTTCAGTACGTGATGCTCCCGTATCGTATCAAAGTTACGAACTAGTAACCGG
GGTGAATGGAAGGCACCGGAATGGTTATCTCTCCAATACCCACTGTCGGAGAAGCGAACACAACACA
```

under Trimmed click on fasta

```
>1-rpoB-S 843 47 644 0.05
CCGATTTCCGCAGCAGTGAAAGAGTTCTTCGGTTCCAGCCAGCTGTCTCAGTTTATGAACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGTC
GTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAAT
CGAAACCCCTGAAGGTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATACGGCTTCCTTGAGACTCCGTATCGTAAA
GTGACCGACGGTGTTGTAACTGACGAAATTCACTACCTGTCTGCTATCGAAGAAGGCAACTACGTTATCGCCCAGGCGAACTCCAACCTGGATGAAG
AAGGCCACTTCGTAGAAGACCTGGTAACTTGCCGTAGCAAAGGCGAATCCAGCTTGTTCAGCCGTGACCAGGTTGACTACATGGACGTATCCACCCA
GCAGGTGGTATCCGTCGGTGCGTCCCTGATCCCGTTCCTGGAACACGATGACGCCACCGTGCATTGATGGGTGCGAACATGCAACGTCAGGCCGTTC
CGACTCTGCGTGCTGATAAGCCGCTGGTTGGTACTGGTATGGAACGTGCTGTTGCCGTTGAC
```