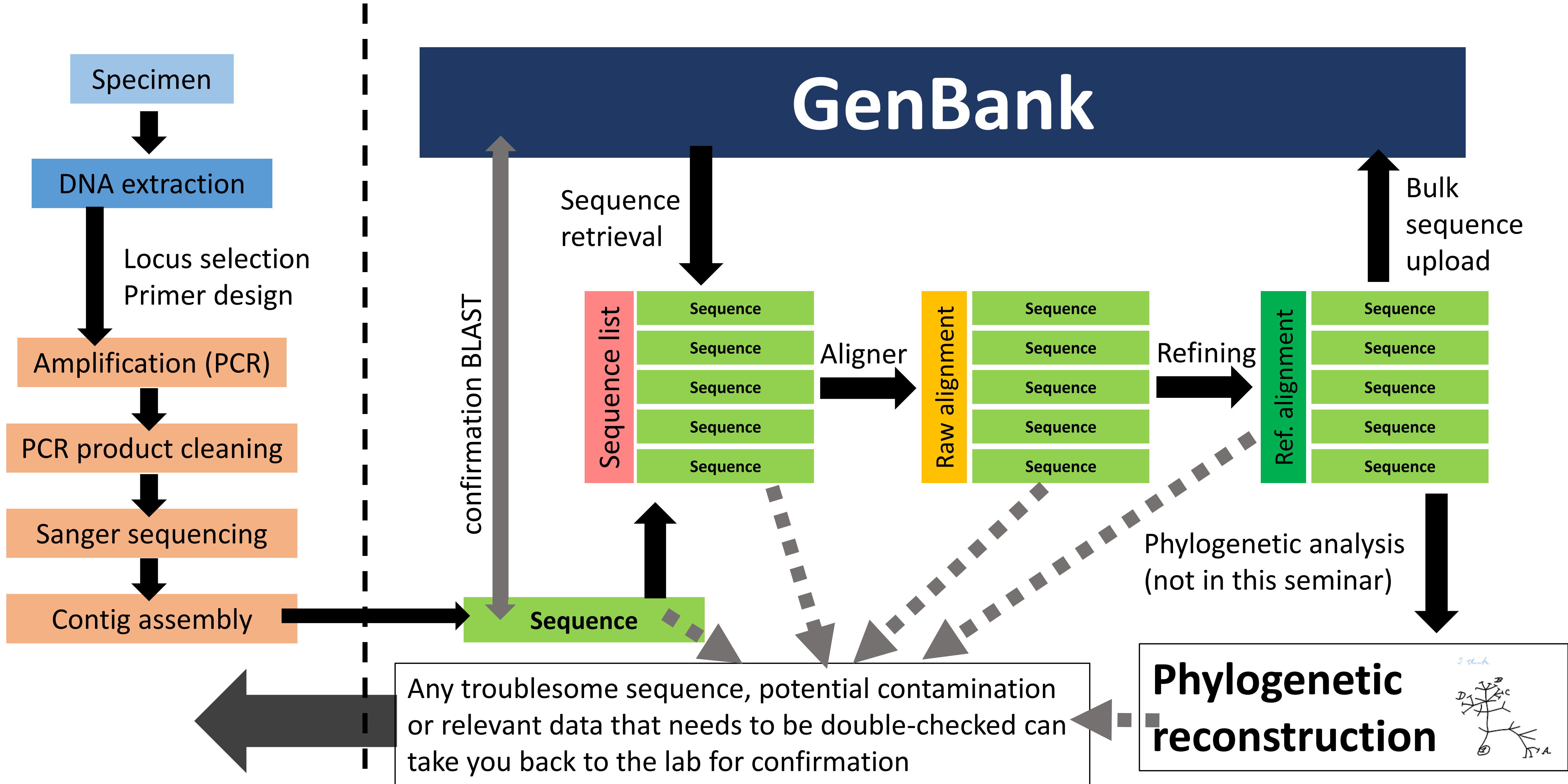
A primer for practical phylogenetic data gathering. Uconn EEB3899-007. Spring 2015

Session 1

DNA extraction, amplification and sequencing

Rafael Medina (rafael.medina.bry@gmail.com)
Yang Liu (yang.liu@uconn.edu)

A road map to phylogenetic data gathering



From tissue to sequences

- **DNA extraction**
- **DNA amplification (PCR)**
- **Validation**
- **PCR product cleaning**
- **Sequencing reaction**
- **Preparation for ABI sequencing**
- **Contig assemblage and sequence exportation**

From tissue to sequences

- **DNA extraction**
- **DNA amplification (PCR)**
- **Validation**
- **PCR product cleaning**
- **Sequencing reaction**
- **Preparation for ABI sequencing**
- **Contig assemblage and sequence exportation**

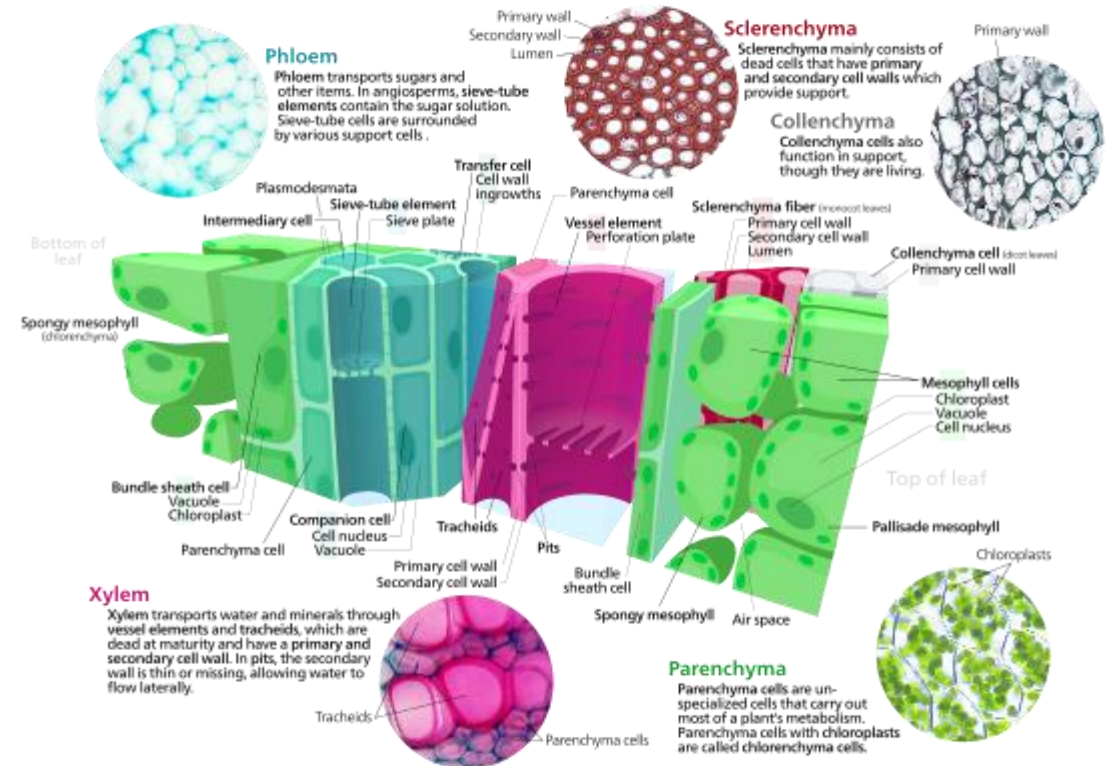
DNA extraction

Typical steps, shared by different kits or protocols:

1. Mechanical homogenization

- Grinding with mortar/pestle and sand/liquid N_2
- Bead beater
- Heat shock

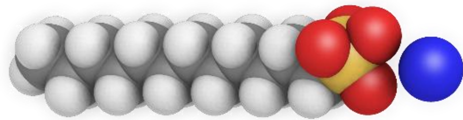
Plant cells: tough nuts to crack!



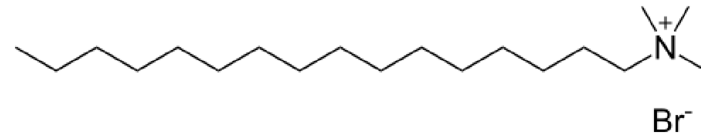
DNA extraction

2. Digestion

Relies on the presence of chaotropic salts, denaturing agents and detergents (often mixed in a “lysis buffer”)



SDS (Sodium Dodecyl Sulfate)



CTAB (cetyltrimethylammonium bromide)

- Performed usually at 65°C
- Longer digestion times may improve results
- RNAse is usually recommended in many protocols

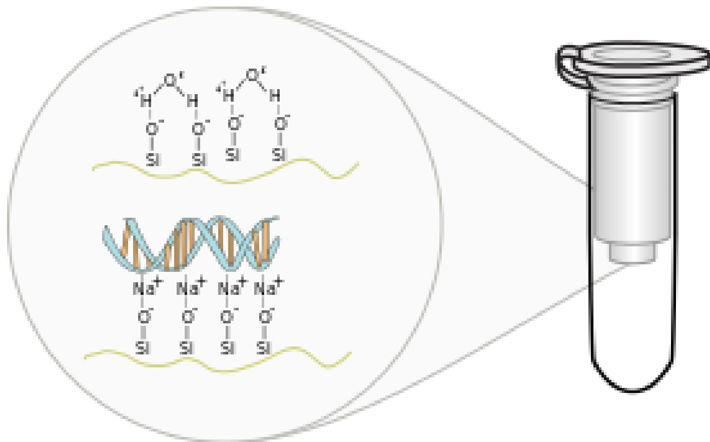
The clean lysate may be transferred after filtering or precipitation of a pellet

DNA extraction

3. Binding and cleaning

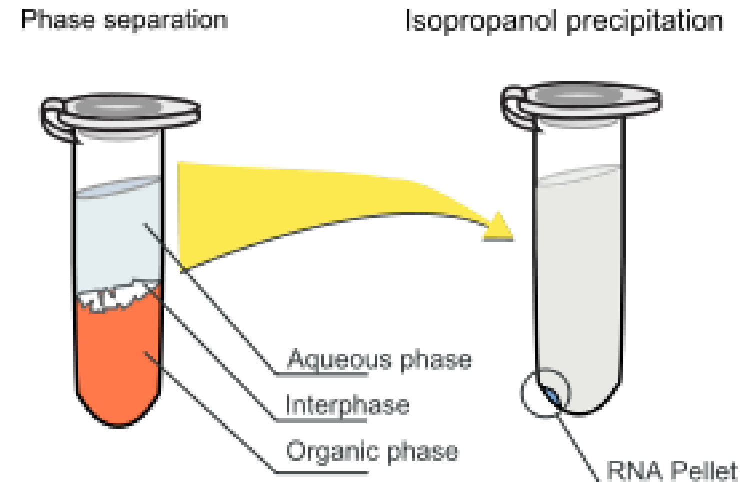
A. Silica membrane kit

- Binding buffer
- DNA stays in the membrane



B. Phenol-chloroform extraction

- Lysate is divided in two phases
- DNA is precipitated with isopropanol



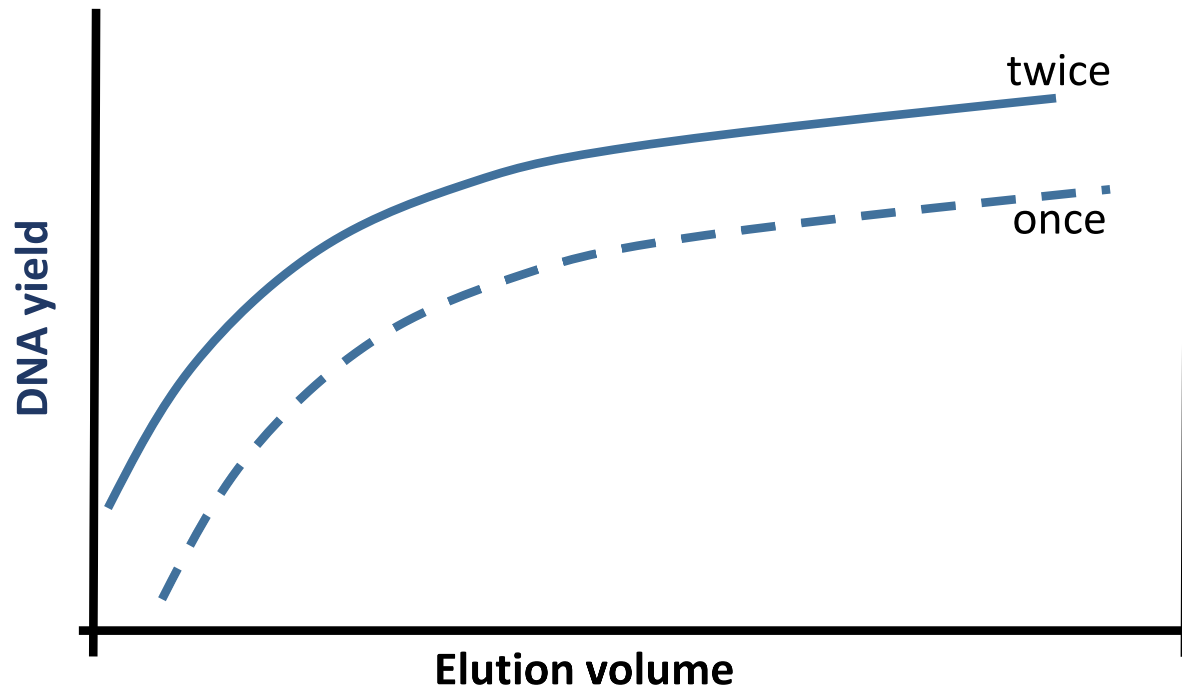
- In both cases, the DNA is washed with ethanol-based buffers
- The ethanol must be evaporated completely before elution

DNA extraction

4. Elution

Usually performed with (warm) Tris-HCl buffer. EDTA based buffers are not recommended

Optimization of results may involve playing with number of elutions, volume and time



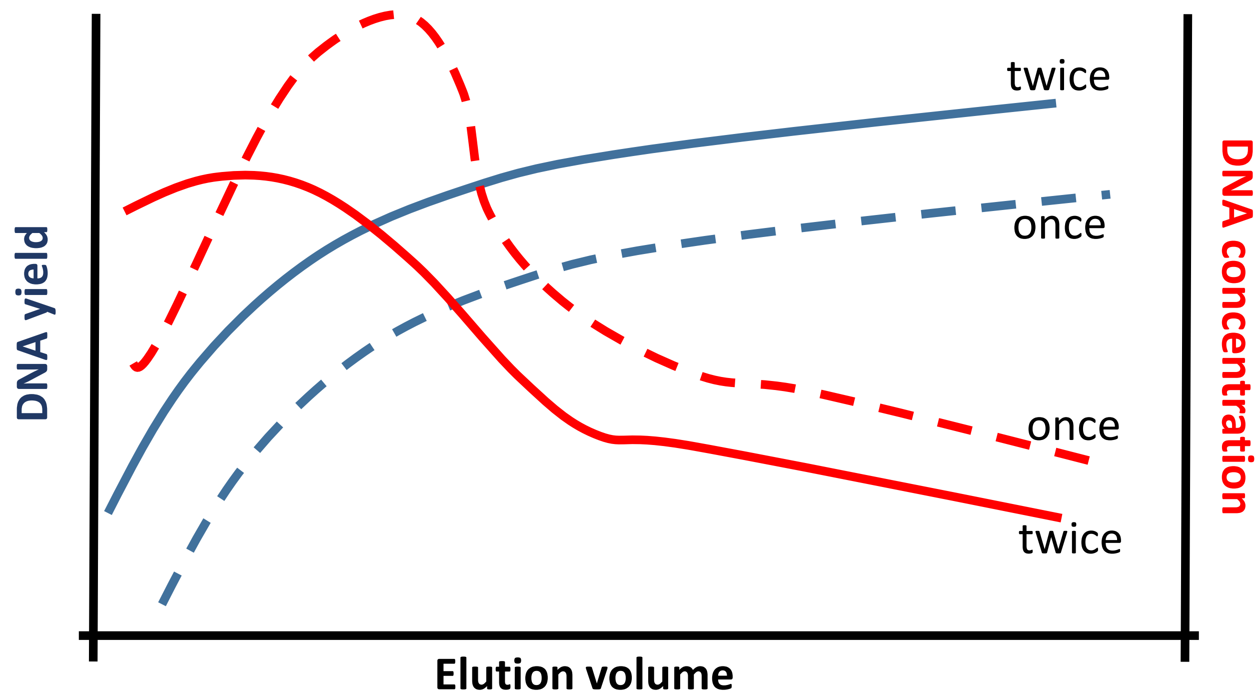
from Macherey-Nagel NucleoSpin

DNA extraction

4. Elution

Usually performed with (warm) Tris-HCl buffer. EDTA based buffers are not recommended

Optimization of results may involve playing with number of elutions, volume and time



from Macherey-Nagel NucleoSpin

DNA extraction

5. Storage

Please, please, PLEASE: label and database your DNA sample so it can be used in the future

Double stranded DNA is very stable in Tris-HCl buffer or even in water

- 20°C for medium term
- 80°C long-term

Quality control (optional)

Nanodrop or Qubit

DNA extraction

5. Storage

Please, please, PLEASE: label and database your DNA sample so it can be used in the future

Double stranded DNA is very stable in Tris-HCl buffer or even in water
- 20°C for medium term
- 80°C long-term

Quality control (optional)

Nanodrop or Qubit

From tissue to sequences

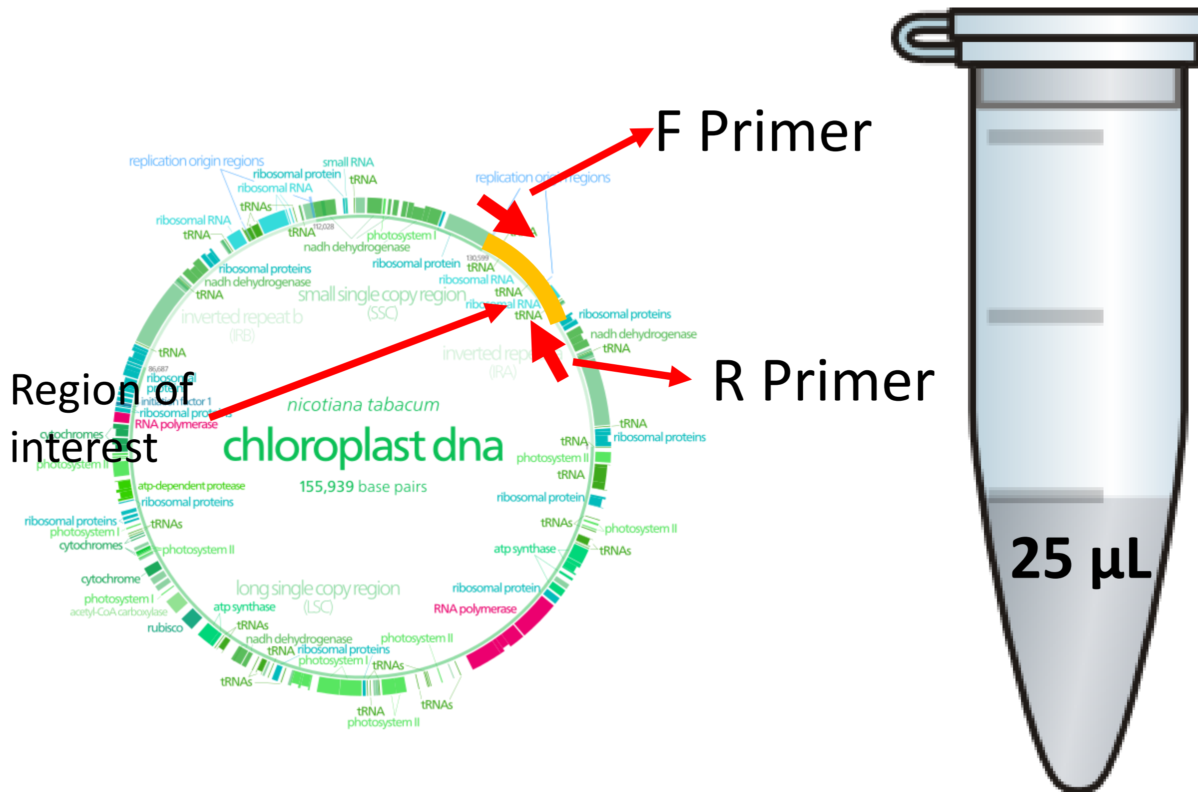
- DNA extraction
- **DNA amplification (PCR)**
- Validation
- PCR product cleaning
- Sequencing reaction
- Preparation for ABI sequencing
- Contig assemblage and sequence exportation

PCR: Polymerase Chain Reaction

(you probably know already about this)

http://youtu.be/iQsu3Kz9NYo

Genomic DNA

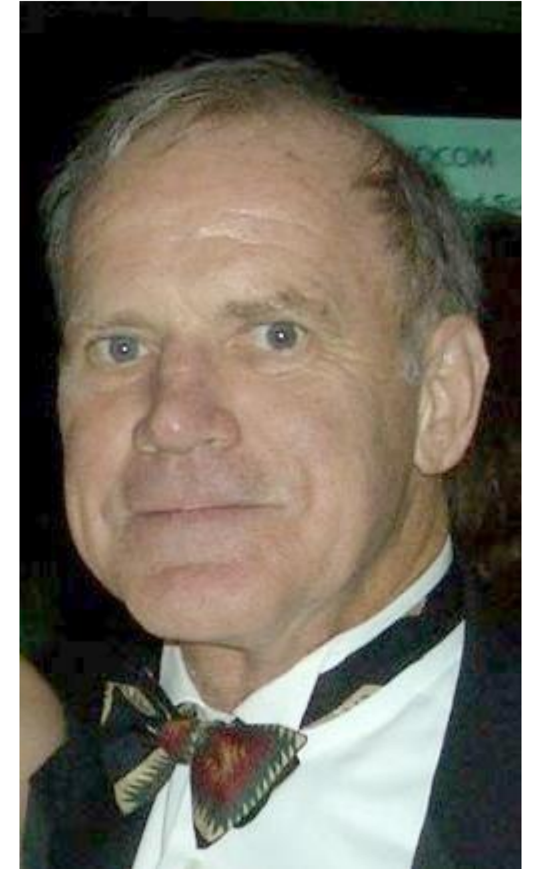


Water

Buffer (with Mg $^{2+}$)

dNTPs

Taq Polymerase

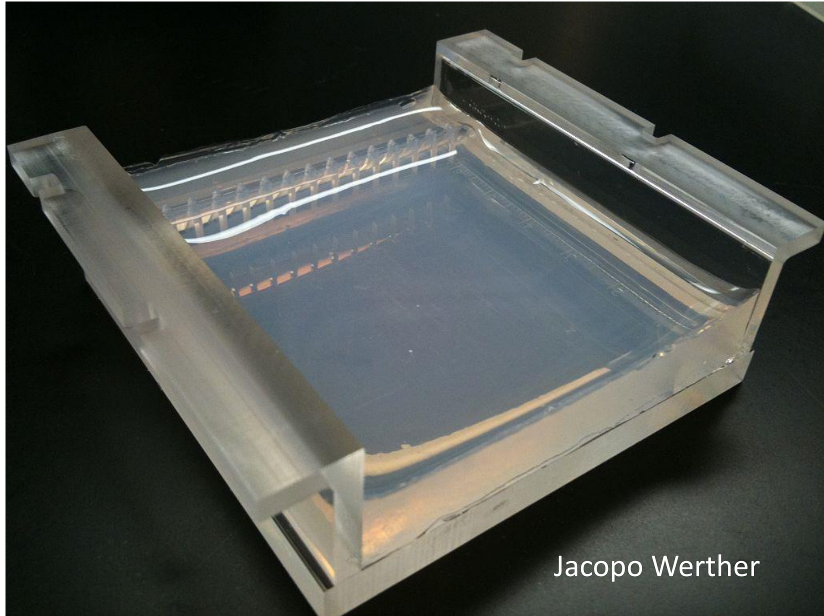


Kary Mullis

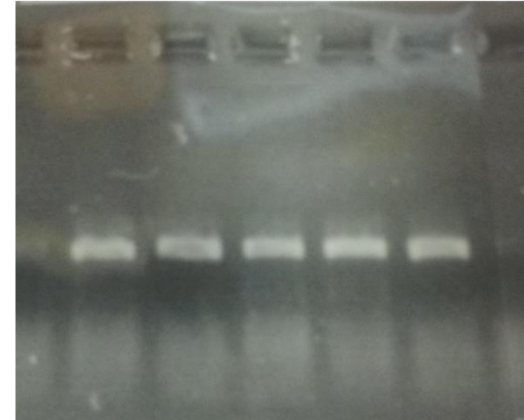
From tissue to sequences

- DNA extraction
- DNA amplification (PCR)
- **Validation**
- PCR product cleaning
- Sequencing reaction
- Preparation for ABI sequencing
- Contig assemblage and sequence exportation

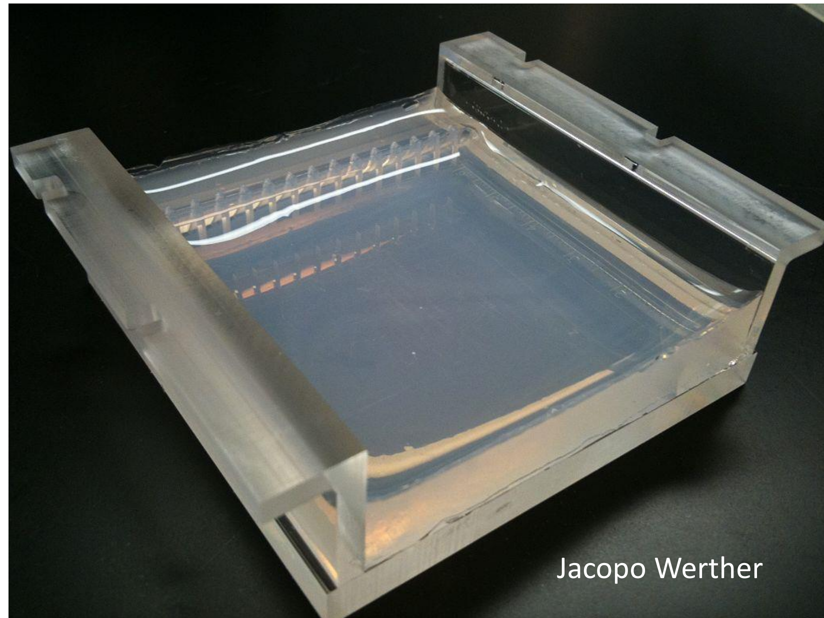
Validation of a PCR reaction: agarose gel electrophoresis



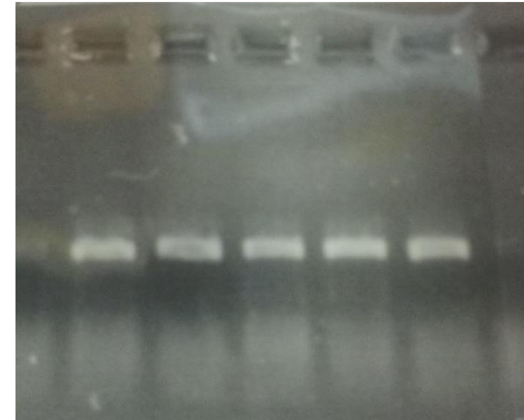
Good, sharp bands



Validation of a PCR reaction: agarose gel electrophoresis



Good, sharp bands



Primer dimer

Double product

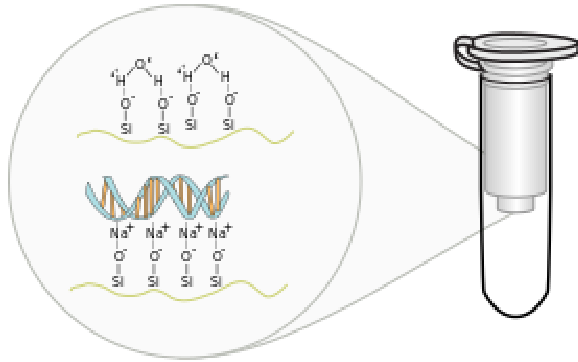
From tissue to sequences

- DNA extraction
- DNA amplification (PCR)
- Validation
- **PCR product cleaning**
- Sequencing reaction
- Preparation for ABI sequencing
- Contig assemblage and sequence exportation

PCR product cleaning

Basically, you need to get rid of primer residues, remaining nucleotides, enzymes, etc

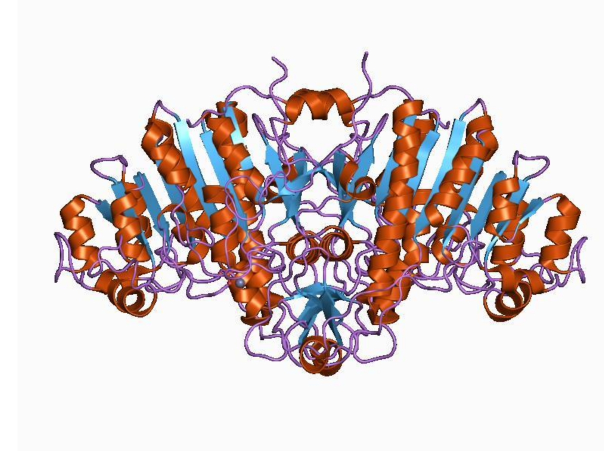
Option 1: commercial kit (silica-membrane) method

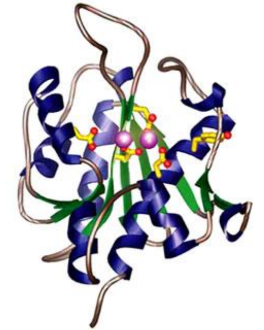


Pro: gives a very clean product
Cons: more expensive and time consuming. Lots of plastic waste

Option 2: enzymatic (Exo-SAP) method

- Shrimp Alkaline Phosphatase (SAP) dephosphorylates remaining dNTPs
- Exonuclease I degrades single-stranded DNA





- Both enzymes are then denatured by heat

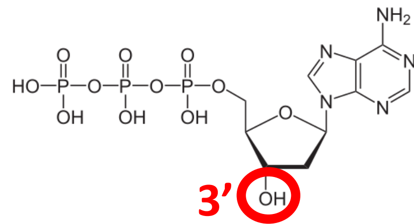
Pros: fast and cheap. Little waste involved
Cons: you do not really “get rid” of the residues. Also: enzymes expire

From tissue to sequences

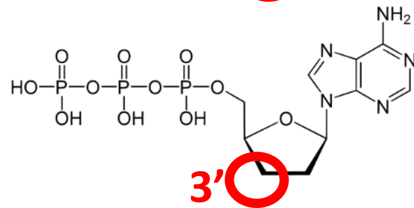
- DNA extraction
- DNA amplification (PCR)
- Validation
- PCR product cleaning
- **Sequencing reaction**
- Preparation for ABI sequencing
- Contig assemblage and sequence exportation

Sequencing reaction

It resembles a normal PCR but
- Using a mix of normal dNTPs and fluorescent labeled ddNTPs



dATP

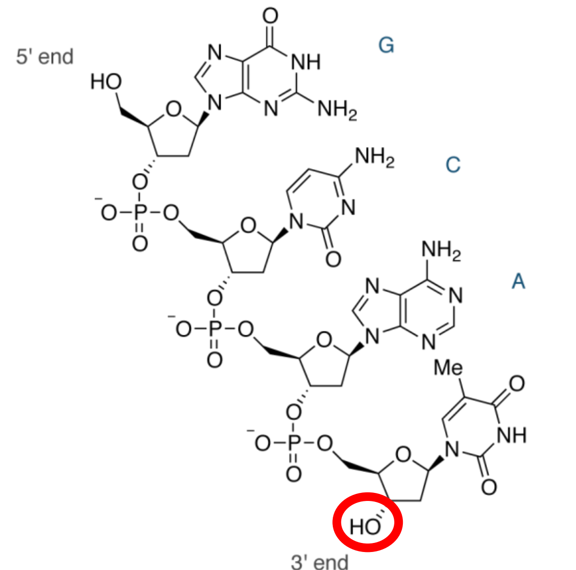


ddATP

A DNA molecule that incorporates a ddnucleotide, cannot grow further

- Using only one primer

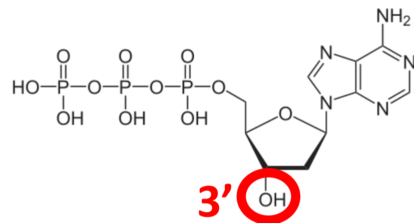
Frederick Sanger



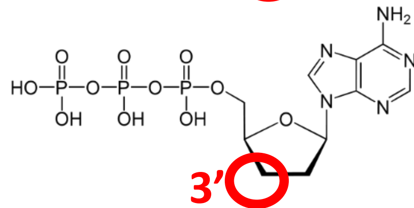
Sequencing reaction

It resembles a normal PCR but

- Using a mix of normal dNTPs and fluorescent labeled ddNTPs



dATP

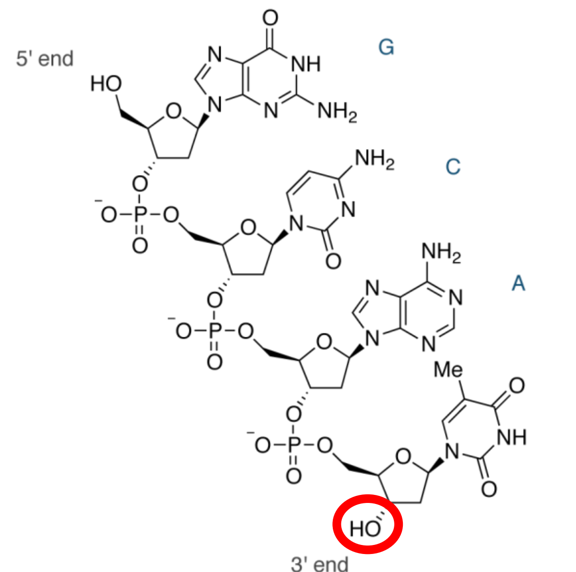


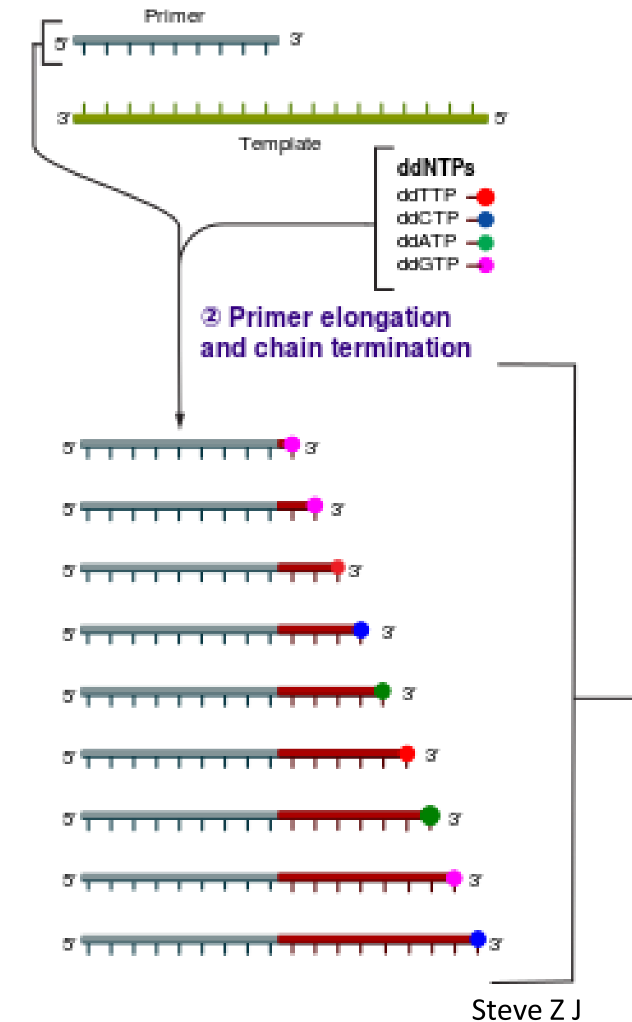
ddATP

A DNA molecule that incorporates a ddnucleotide, cannot grow further

- Using only one primer

Frederick Sanger





Steve Z J

This results in a population of fragments with all the possible sizes and with a fluorescent 3’ terminal nucleotide

From tissue to sequences

- DNA extraction
- DNA amplification (PCR)
- Validation
- PCR product cleaning
- Sequencing reaction
- **Preparation for ABI sequencing**
- Contig assemblage and sequence exportation

Preparation for ABI sequencing

Plan in advance the use of the ABI (bunches of 16 sequences), book a time slot on the whiteboard

Filter your samples through a Sephadex G-50 column into the sequencing plate

Have ready a spreadsheet with the information of each of your samples and the well it will take in the sequencing plate. All that information should also be in you lab book

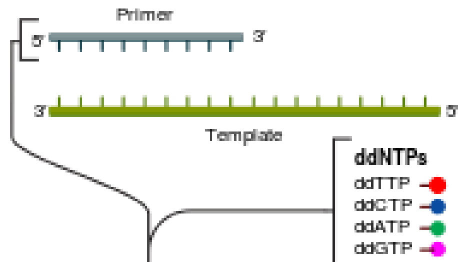
Container Name	Description	ContainerType	AppType	Owner	Operator	
Medinaseq19_23Jun2014		96-Well	Regular	Medina	Medina	
AppServer	AppInstance					
SequencingAnalysis						
Well	Sample Name	Comment	Priority	Results Group 1	Instrument Protocol 1	Analysis Protocol 1
A01	Physcomitrella_magdalenae_3828_atpbrbclF		100	GoffinetLab_sequencing	Seq50-POP7-E(BD1.1)	Seq_E
B01	Physcomitrella_magdalenae_3828_psbAtrnHF		100	GoffinetLab_sequencing	Seq50-POP7-E(BD1.1)	Seq_E
C01	Physcomitrium_bukobense_3829_atpbrbclF		100	GoffinetLab_sequencing	Seq50-POP7-E(BD1.1)	Seq_E

Taxon name, DNA ID, Region, primer

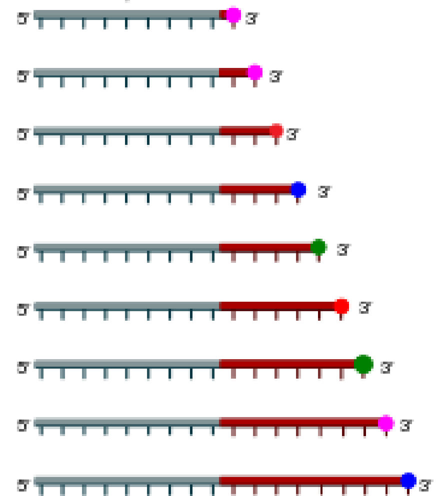
Destination folder

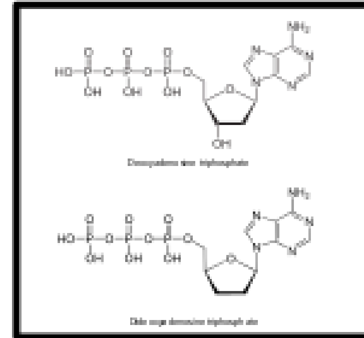
ABI Sequencing

① Reaction mixture
- Primer and DNA template
- DNA polymerase
- ddNTPs with flourochromes
- dNTPs (dATP, dCTP, dGTP, and dTTP)

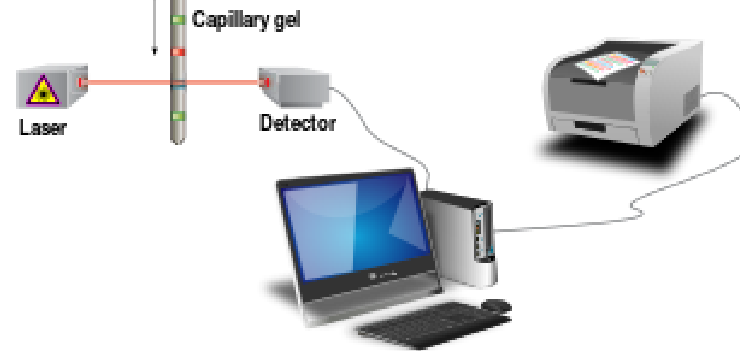


② Primer elongation and chain termination

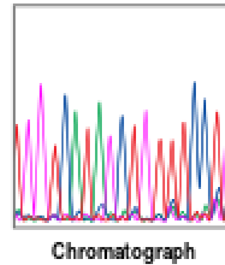




③ Capillary gel electrophoresis separation of DNA fragments

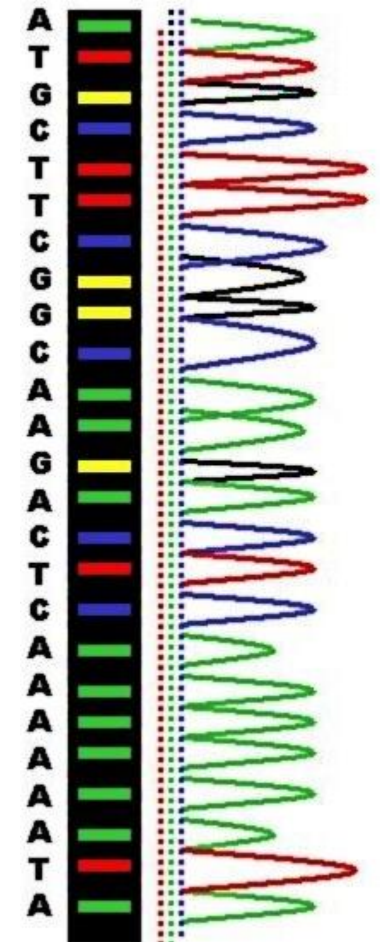


Steve Z J



Chromatograph

④ Laser detection of flourochromes and computational sequence analysis



Abizar Lakdawalla

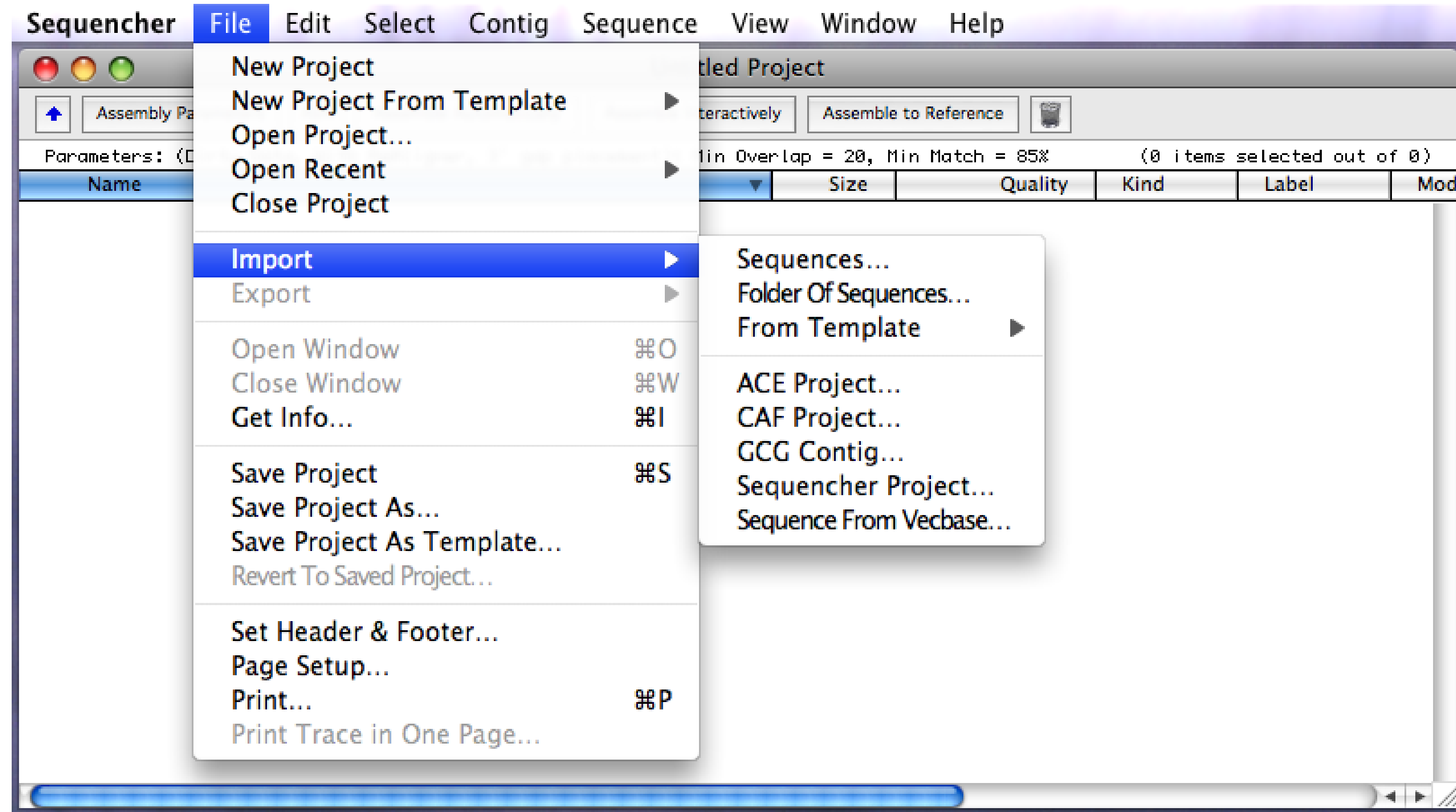
Output of the ABI: a chromatogram in .ab1 format

From tissue to sequences

- DNA extraction
- DNA amplification (PCR)
- Validation
- PCR product cleaning
- Sequencing reaction
- Preparation for ABI sequencing
- **Contig assemblage and sequence exportation**

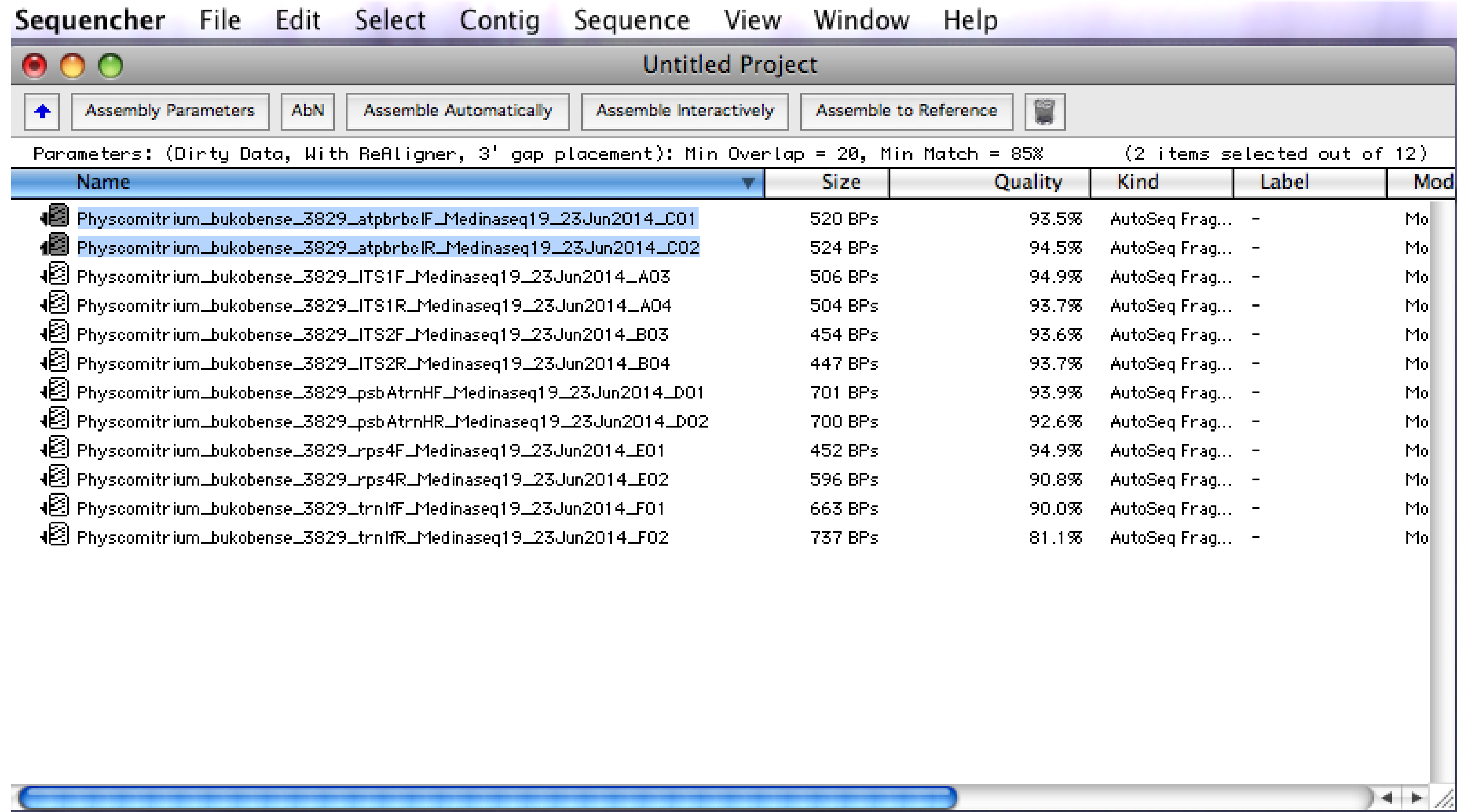
Contig assemblage

Import your ab1 files into a Sequencher (Geneious, etc) project



Contig assemblage

Import your ab1 files into a Sequencher (Geneious, etc) project



Contig assemblage

Assemble automatically your sequence pair

Sequencher File Edit Select Contig Sequence View Window Help

Untitled Project

Assembly Parameters | AbN | Assemble Automatically | Assemble Interactively | Assemble to Reference

Parameters: (Dirty Data, With ReAligner, 3' gap placement): Min Overlap = 20, Min Match = 85% (2 items selected out of 12)

Name	Size	Quality	Kind	Label	Mod
Contig[0002]	548 BPs		Contig of 2	-	We
Physcomitrium_bukobense_3829_ITS1F_Medinaseq19_23Jun2014_A03	506 BPs	94.9%	AutoSeq Frag...	-	Mo
Physcomitrium_bukobense_3829_ITS1R_Medinaseq19_23Jun2014_A04	504 BPs	93.7%	AutoSeq Frag...	-	Mo
Physcomitrium_bukobense_3829_ITS2F_Medinaseq19_23Jun2014_B03	454 BPs	93.6%	AutoSeq Frag...	-	Mo
Physcomitrium_bukobense_3829_ITS2R_Medinaseq19_23Jun2014_B04	447 BPs	93.7%	AutoSeq Frag...	-	Mo
Physcomitrium_bukobense_3829_psbAtrnHF_Medinaseq19_23Jun2014_D01	701 BPs	93.9%	AutoSeq Frag...	-	Mo
Physcomitrium_bukobense_3829_psbAtrnHR_Medinaseq19_23Jun2014_D02	700 BPs	92.6%	AutoSeq Frag...	-	Mo
Physcomitrium_bukobense_3829_rps4F_Medinaseq19_23Jun2014_E01	452 BPs	94.9%	AutoSeq Frag...	-	Mo
Physcomitrium_bukobense_3829_rps4R_Medinaseq19_23Jun2014_E02	596 BPs	90.8%	AutoSeq Frag...	-	Mo
Physcomitrium_bukobense_3829_trnlfF_Medinaseq19_23Jun2014_F01	663 BPs	90.0%	AutoSeq Frag...	-	Mo
Physcomitrium_bukobense_3829_trnlfR_Medinaseq19_23Jun2014_F02	737 BPs	81.1%	AutoSeq Frag...	-	Mo

Assembly Completed

Time Elapsed: 00:00:00
Items Selected: 2
Comparisons Performed: 2

Assemble by Name: Off
Number of Contigs: 1
Number of Fragments: 0

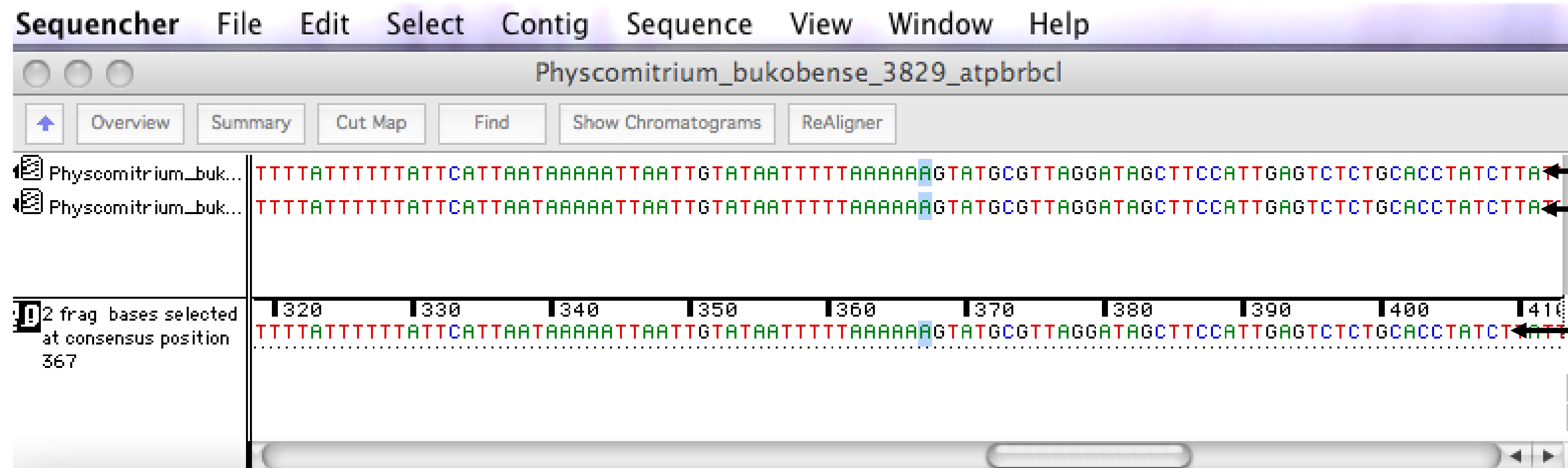
Close

Contig assemblage

Rename the contig with the relevant information

Contig assemblage

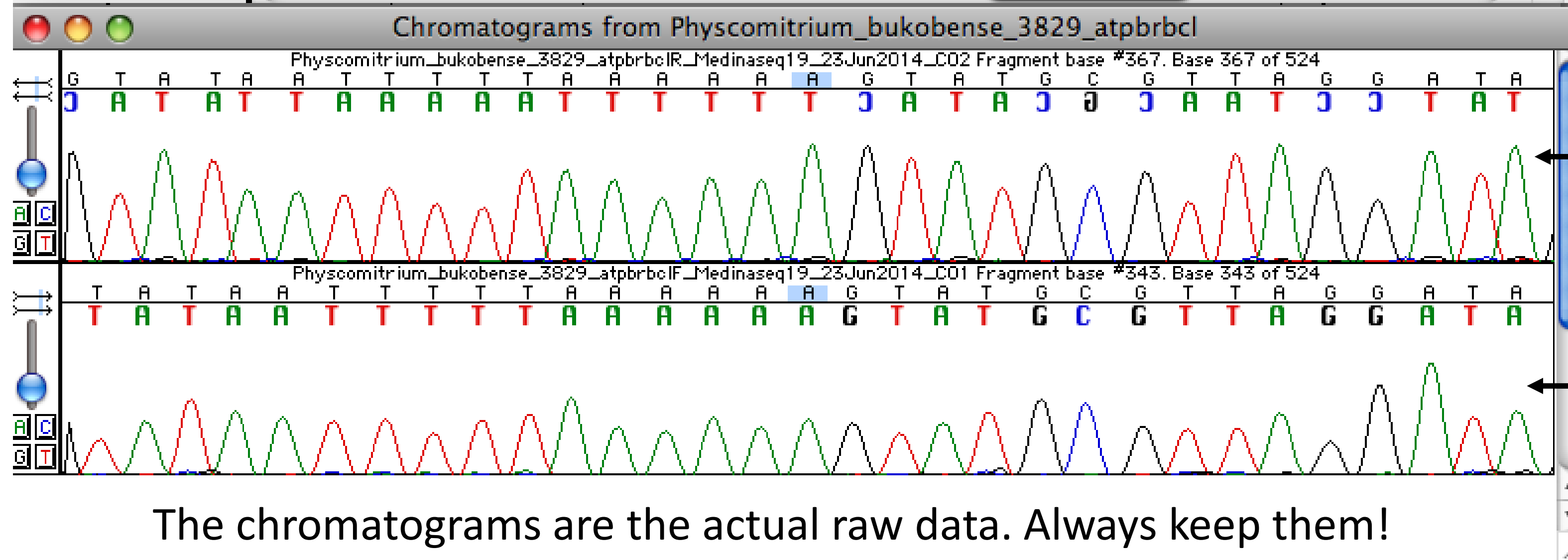
Contig edition window



Interpretation of chromatograms

Consensus sequence (This will become your sequence)

Chromatogram views



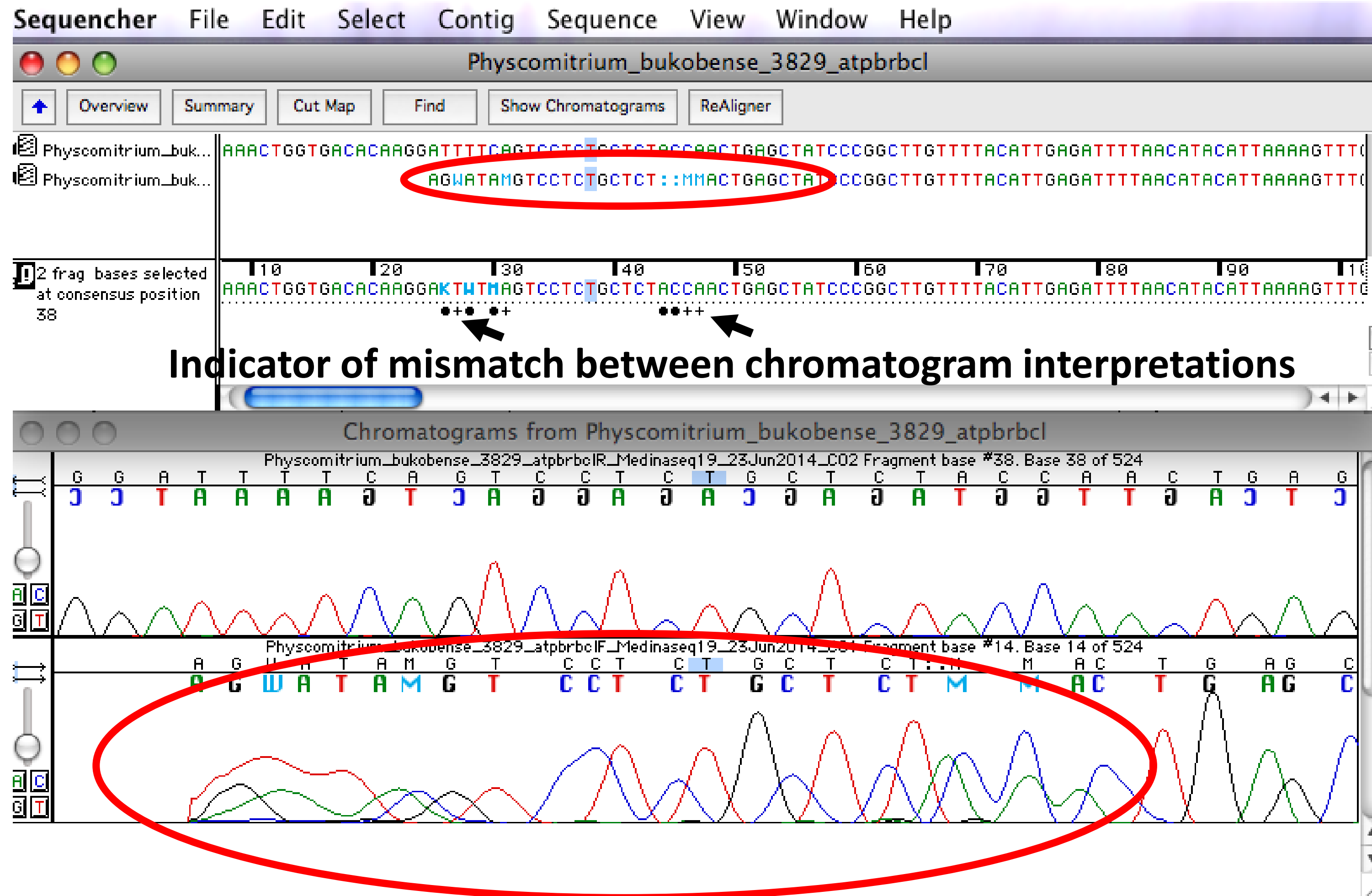
R chromatogram

F chromatogram

The chromatograms are the actual raw data. Always keep them!

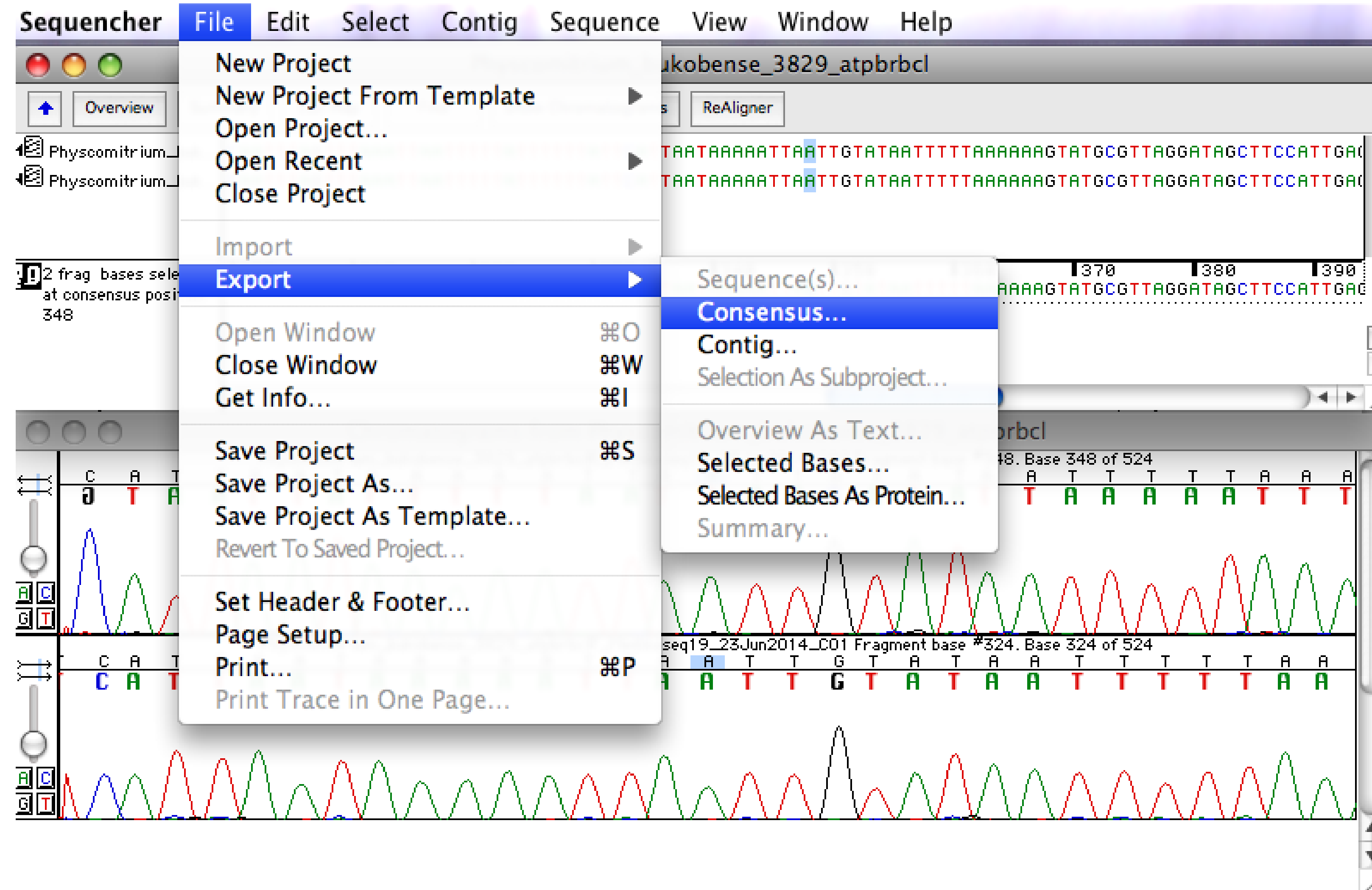
Contig assemblage

The chromatograms are usually "dirty" at the extremes



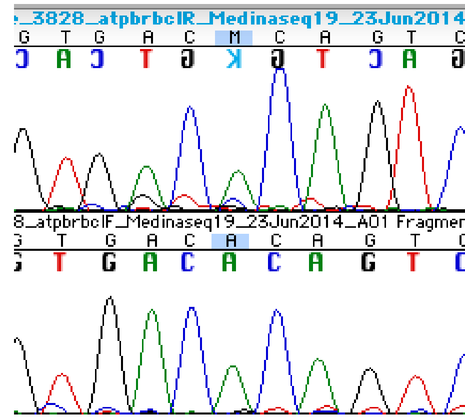
Contig assemblage

Export the consensus sequence as a FASTA (.fas) file when ready

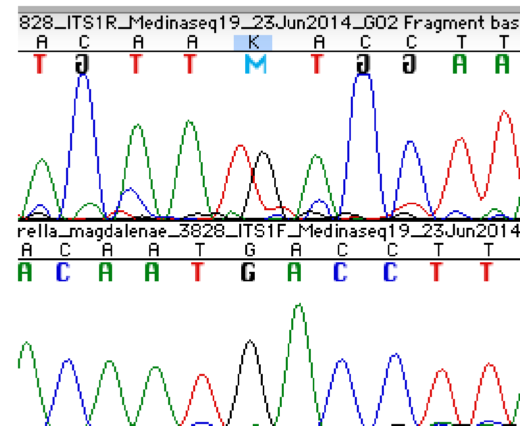


Contig assemblage

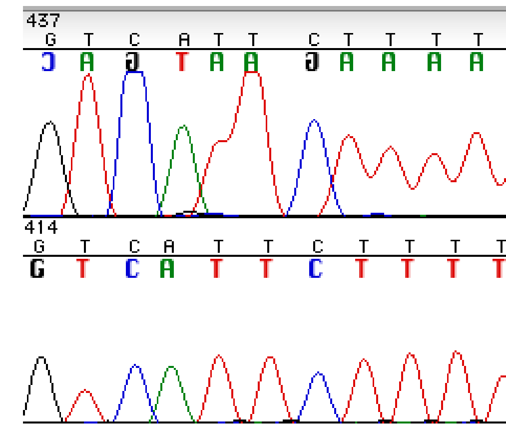
What to expect when editing sequences (best case scenarios)



Faint signal



Merged peaks



Lack of definition

Problematic, noisy sequences can (sometimes) be edited (as we will see in the next session)

Always keep a backup of the raw chromatogram files, the Sequencher/Geneious project and your fasta sequences