



# *Blastocystis*

under One Health

---

## Protocol Handbook



*Blastocystis*  
Identification



from Microscopy  
to Phylogeny

---



Funded by  
the European Union

2<sup>nd</sup> Training School of *Blastocystis* under One Health CA21105 COST Action:

## ***Blastocystis* Identification: From Microscopy to Phylogeny**

University of Kent, Canterbury, UK  
30<sup>th</sup> of August – 1<sup>st</sup> of September 2023

Lectures and computer practicals will take place in **Stacey Lecture Theatre 2**  
Laboratory Practical activities will take place in **Teaching Lab 3 in the Stacey Building**  
Coffee and lunch breaks will take place in **Howard Rogers Room in the Ingram Building**

**You will work in groups of two throughout this workshop. Your group number is on the bench.**

**All practicals require the use of gloves.**

### **Day 1 – 30<sup>th</sup> of August, 2023**

*Instructors: Dr. Anastasios Tsaousis, Prof. Funda Dođruman-Al, Prof. Ayman A. El-Badry, Dr. Eleni Gentekaki*

08:30 – 09:00	Registration – Howard Rogers Room in Ingram Building
09:00 – 09:55	Lecture - Introduction to the Training School and Activities - <i>Dr. Anastasios Tsaousis</i>
09:15 – 09:35	Lecture - Microscopy and identification- <i>Prof. Ayman A. El-Badry</i>
09:35 – 09:55	Lecture – Sample processing- <i>Prof. Funda Dođruman-Al</i>
10:00 – 10:30	Coffee break
10:30 – 13:00	Practical – DNA extraction and quantification- <i>Dr. Anastasios Tsaousis, Dr. Eleni Gentekaki</i>
13:00 – 14:00	Lunch break
14:00 – 15:00	Practical – PCR with universal eukaryotic primers- <i>Dr. Anastasios Tsaousis, Dr. Eleni Gentekaki</i>
15:00 – 15:30	Practical – Microscopy- <i>Prof. Funda Dođruman-Al, Prof. Ayman A. El-Badry</i>
15:30 – 16:00	Tea break
16:00 – 17:00	Practical – Microscopy- <i>Prof. Funda Dođruman-Al, Prof. Ayman A. El-Badry</i>
17:00 – 17:30	Practical – PCR 2 with barcoding primers <b>and</b> direct PCR- <i>Dr. Anastasios Tsaousis, Dr. Eleni Gentekaki</i>
18:00	End of the day

## 1. DNA extraction

On your bench you will find two tubes:

A 1.5 ml tube labelled as “Jones media” containing *Blastocystis* growing in Jones media (200-250  $\mu$ l)

A 1.5 ml tube labelled as “stool” containing non-human stool sample (200-250 mg)

You will extract DNA from the contents of both tubes using the PureLink Microbiome DNA purification kit from Invitrogen.

### *Protocol*

#### 1. Prepare the lysate

Take two Bead Tubes

- a. Tube 1: Add 200-250 mg of stool sample + 650  $\mu$ l S1—Lysis Buffer  
Tube 2: Add 200-250  $\mu$ l of *Blastocystis* in Jones media + 650  $\mu$ l S1—Lysis Buffer
- b. Vortex for 5-10 seconds
- c. Add 100  $\mu$ l of S2—Lysis Enhancer
- d. Vortex for 5-10 seconds
- e. Incubate at 65°C for 10 minutes in a water bath
- f. Vortex tubes at maximum speed for 10 minutes to homogenize
- g. Centrifuge at 14,000  $\times$  g for 2 minutes
- h. Transfer up to 400  $\mu$ l of the supernatant to a clean microcentrifuge tube
- i. Add 250  $\mu$ l of S3—Cleanup Buffer
- j. Vortex immediately to ensure even dispersion of S3—Cleanup Buffer and uniform precipitation of inhibitors
- k. Centrifuge at 14,000  $\times$  g for 2 minutes
- l. Transfer up to 500  $\mu$ l of the supernatant to a clean microcentrifuge tube, avoiding the pellet and any debris.

#### 2. Bind the DNA to the column

- a. Add 900  $\mu$ l of S4—Binding Buffer
- b. Vortex for 5-10 seconds

- c. Load 700  $\mu\text{l}$  of the sample mixture onto a spin column-tube
- d. Incubate at room temperature for 2 minutes
- e. Centrifuge at  $14,000 \times g$  for 1 minute
- f. Discard the flow-through and repeat steps 2c-2e with the remaining sample mixture

*Note: Ensure that the entire sample mixture has passed into the collection tube by inspecting the column. If sample remains in the column, centrifuge again at  $14,000 \times g$  for 1 minute*

### **3. Wash and elute the DNA**

- a. Place the spin column in a clean collection tube
- b. Add 500  $\mu\text{l}$  of S5—Wash Buffer
- c. Incubate at room temperature for 2 minutes
- d. Centrifuge the spin column-tube at  $14,000 \times g$  for 1 minute
- e. Discard the flow-through
- f. Centrifuge the spin column-tube at  $14,000 \times g$  for another 30 seconds. (The second centrifugation optimizes the removal of S5—Wash Buffer, which could interfere with downstream applications)
- g. Place the spin column in a clean tube, add 100  $\mu\text{l}$  of S6—Elution Buffer
- h. Incubate at room temperature for 1 minute
- i. Centrifuge the spin column-tube at  $14,000 \times g$  for 1 minute, then discard the column  
The purified DNA is in the tube

## **2. Quantification of DNA**

For the quantification of DNA, you will use BioDrop

## **3. Polymerase chain reaction (PCR)**

In this workshop, you will do a nested PCR and a direct PCR. The nested PCR requires two reactions, one with universal eukaryotic primers (RD5 and RD3) and another with the barcoding primers (BsRD5F and BhRDr9R). The direct PCR is a single reaction with the barcoding primers. The direct PCR is the most commonly used approach, as it is faster and reduces false positives. The product size for the second reaction of the nested PCR and the direct PCR is  $\sim 600\text{bp}$ .

Primers sequences:

BhRDr 5'-GAGCTTTTTAACTGCAACAACG-3'  
 RD5F 5'-ATCTGGTTGATCCTGCCAGT-3'  
 RD3 5'-GGGATCCTGATCCTTCCGCAGGTTACCTAC-3'  
 RD5 5'-GGAAGCTTATCTGGTTGATCCTGCCAGTA-3'

Important to know before starting:

- Keep all primers and PCR reagents on ice.
- Wipe all pipettes with IMS before every PCR (also make sure you don't touch the pipette in any of the tubes as this could lead to contamination).
- For the negative control, replace 1 µl of DNA with 1 µl of milliQ sterile water (i.e. no DNA template in this control). For the positive control, a known positive sample will be provided.
- When you make the Mastermix, do **not** add the DNA. Add your sample DNA to each individual PCR tube/qPCR wells once the mastermix has been aliquoted.
- The volume included below includes the volume required for one sample. You will then need to calculate the total volume by considering the number of samples you have, including a positive and negative control, as well as an extra reaction mix. Let's say, for example, you have 2 samples, along with a positive and negative control, and an additional spare reaction mix, This implies that you will need to multiply the volume for one sample by the total number of reactions, which in this case would be 5.

*Protocol nested PCR, reaction 1*

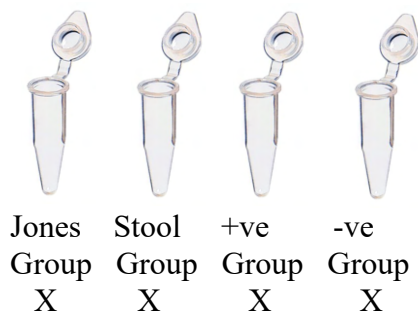
Primers used: RD5 + RD3

You will need to set up four tubes. Two tubes will contain the DNA from Jones and stool that you extracted earlier, one tube will contain water instead of DNA (negative control) and one will contain DNA from a *Blastocystis* positive sample (positive control). The DNA of the positive control comes from a pure culture and has been extracted for you.

The total volume of each reaction is 50 µl. Each reaction should contain:

- 25 µl PCRBIO Taq Mix Red
- 2 µl RD3
- 2 µl RD5
- 19 µl water
- 2 µl DNA

Label your tubes as follows:



X=your group number

**Reminder:** always make the master mix for an additional reaction than what you plan (e.g. if you have 4 samples, make master mix for 5 samples)

a. Master mix preparation: This is prepared by mixing the reagents that are common to all tubes so that you save time and reduce error. In a sterile 1.5 ml tube mix the following:

Forward primer RD5:	5 x 2 $\mu$ l = 10 $\mu$ l
Reverse primer RD3:	5 x 2 $\mu$ l = 10 $\mu$ l
PCRBIO Taq Mix Red:	5 x 25 $\mu$ l = 125 $\mu$ l
Water:	5 x 19 $\mu$ l = 95 $\mu$ l
Final:	240 (equals 48 $\mu$ l per PCR tube)

b. Mix well by pipetting up and down gently

c. Distribute 48  $\mu$ l of the master mix in each tube

d. Add 2  $\mu$ l of water in the –ve control tube and close it

e. Add 2  $\mu$ l of DNA (as shown in figure above) and close the tubes

f. Finally, add 2  $\mu$ l of +ve control DNA in +ve control tube and close it

g. Insert tubes in PCR machine

h. Run program. This will be pre-programmed for you. The PCR protocol is below for your reference

- 95°C for 5 minutes
  - 94°C for 1 minute
  - 59°C for 1 minute
  - 72°C for 1 minute
  - 72°C for 10 minutes
  - Hold at 12°C
- } x30

Fragment size for 1st PCR: 1780 bp

#### 4. Microscopy

In this part of the workshop you will observe permanent slides stained with trichrome stain.

##### *Trichrome stain protocol*

The trichrome stain has already been done for you. The protocol for the stain is provided below for future reference. Observe the slides of stained *Blastocystis*

#### Reagents and preparation of the solutions

##### *Trichrome stain*

Chromotrope 2R.....	6 g
Light green SF.....	3 g

Phosphotungstic acid.....7 g  
 Acetic acid (glacial).....10 ml  
 Distilled water.....1000 ml

1. Add 10 ml of glacial acetic acid to 6 g of chromotrope 2R, 3 g of light green SF and 7 g of phosphotungstic acid in a clean flask.
2. Swirl to mix and leave to stand for 30 minutes.
3. Add 1000 ml of distilled water and mix thoroughly; the stain should be deep purple.
4. Store in a glass-stoppered bottle; the stain is stable and is used undiluted.

***D'Antoni's Iodine***

Iodine crystals (I<sub>2</sub>) .....15.0 g  
 Potassium Iodide (KI).....10.0 g  
 Distilled Water .....1000.0 ml

1. Dissolve iodine crystals in distilled water.
2. Add potassium iodide to this solution then shake until a dark reddish brown strong-tea colour is obtained and saturated.
3. Filter and keep this solution in the dark brown glass bottle

***Schaudinn's Fixative***

*Mercuric Chloride, Saturated Aqueous Solution*

Mercuric chloride (HgCl<sub>2</sub>) ..... 110 g  
 Distilled water.....1,000 ml

1. Use a beaker as a water bath; boil (use a hood if available) until the mercuric chloride is dissolved; let stand for several hours until crystals form.

*Schaudinn's Fixative (Stock Solution mercuric chloride base)*

Mercuric chloride, saturated aqueous solution..... 600 ml  
 Ethyl alcohol, 95%..... 300 ml  
 Glycerine..... 15 ml

1. Mix all components and immediately before use, add 5 ml of glacial acetic acid per 100 ml of stock solution.

*Copper Sulphate, Saturated Aqueous Solution*

Copper Sulphate (CuSO<sub>4</sub>.5H<sub>2</sub>O)..... 20 g  
 Distilled water..... 1,000 ml

1. Use a beaker as a water bath; boil (use a hood if available) until the copper sulphate is dissolved; let stand for several hours until cold and keep in a bottle.

*Schaudinn's Fixative (Stock Solution copper sulphate base)*

Copper Sulphate saturated aqueous solution..... 600 ml  
 95% Ethyl alcohol.....300 ml  
 Glycerine.....15 ml

1. Mix all components and immediately before use, add 5 ml of glacial acetic acid per 100 ml of stock solution.

### ***Carbol Xylene***

Xylene.....90 ml  
Phenol (carbolyc acid) crystal.....30 ml

1. Phenol crystals are melted in a bain-marie (water bath) to 30 ml and combined with xylene.

### ***90% Ethanol, Acidified***

90% Ethanol.....99.5 ml  
Glacial acetic acid..... 0.5 ml

1. Mix components in a glass container.

### **Xylene**

#### **Procedure notes for trichrome staining**

-Depending on the volume of slides stained, staining solutions will have to be changed on an as-needed basis.

-If the xylene becomes cloudy or there is an accumulation of water in the bottom of the staining dish containing xylene, discard the old reagents, clean the dishes, dry them thoroughly, and replace them with fresh xylene.

-All staining dishes should be covered to prevent evaporation of reagents (screw-cap Coplin jars or glass lids).

-The single most important step in the preparation of a well-stained faecal smear is good fixation

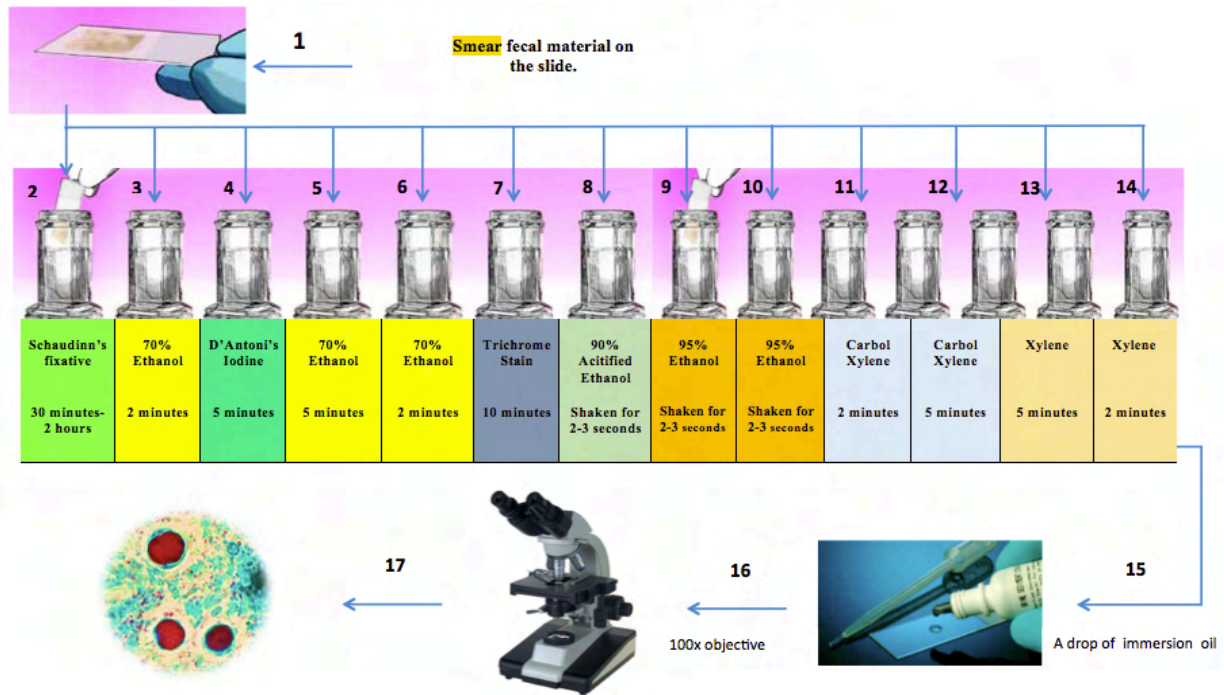
-Slides should always be drained between solutions. Touch the end of the slide to a paper towel for 2 seconds to remove excess fluid before proceeding to the next step. This will maintain the staining solutions for a longer period of time. The slide can also be touched against the staining dish to drain off excess fluid before moving on to the next dish.

-When the smear is thoroughly fixed and the staining procedure is performed correctly, the cytoplasm of protozoan trophozoites is blue-green, sometimes with a tinge of purple. Cysts tend to be slightly more purple. Nuclei and inclusions (chromatoidal bars, RBCs, bacteria, and Charcot-Leyden crystals) are red, sometimes tinged with purple. The background material usually stains green, providing a nice colour contrast with the protozoa.

#### **References**

Garcia LS. Diagnostic Medical Parasitology 6th edition. ASM Press 2016, USA Bench aids for the diagnosis of intestinal parasites, second edition. Geneva: World Health Organization;2019



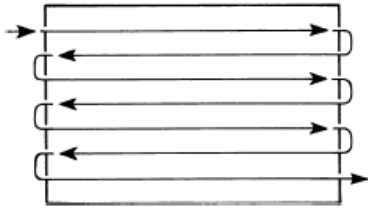


### Direct stool wet smear protocol

1. Take a clean and grease-free slide.
2. Place one drop of physiological saline on the left side of the slide and one drop of iodine on the right side of the same slide.
3. Take a small amount of stool specimen (~2 mg or rice grain size) with the plastic loop and place it into the saline solution drop, repeat the same step for the iodine drop by using another stick.
  - \*If the stool is solid or shaped/soft, a sample is taken by dipping the inside of the stool several times from different places with a stick (use separate sticks for saline and iodine sides).
  - \*If the stool is watery, a sample is taken from the bottom of the sample container with a Pasteur pipette, a drop of the sample is dropped into the saline and iodine and homogenized with the help of a swab. In very liquid samples, a coverslip is closed directly on the sample drop without using saline.
4. Thoroughly emulsify the stool in the saline and iodine preparations then place a coverslip on each suspension without creating bubbles.

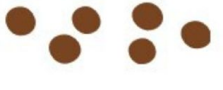






*Note: If a fresh stool specimen is received and if blood and mucus are present, the specimen should be examined as a direct mount making sure to sample the bloody areas.*

5. Systematically scan the entire coverslip with the low power objective (10x) and low light intensity.
6. If you encounter any suspicious objects, examine the specimen with a higher magnification (40x).
7. At least one-third to one-half of the coverslip should be examined under higher magnification (40x objective), even if nothing suspicious has been seen.



**Fig.-**Suggested method of scanning direct wet film preparation with a 10× objective. Note that the entire coverslip preparation should be examined before indicating the examination is negative. -

*Note: Bristol Stool Scale is a diagnostic medical tool designed to classify the form of human faeces into seven categories (Lewis SJ&Heaton KW, 1997)*

Type 1		Separate hard lumps, like nuts (hard to pass)
Type 2		Sausage-shaped but lumpy
Type 3		Like a sausage but with cracks on its surface
Type 4		Like a sausage or snake, smooth and soft
Type 5		Soft blobs with clear-cut edges
Type 6		Fluffy pieces with ragged edges, a mushy stool
Type 7		Watery, no solid pieces, entirely liquid

#### References

Garcia LS. Diagnostic Medical Parasitology 6th edition. ASM Press 2016, USA  
 Bench aids for the diagnosis of intestinal parasites, second edition. Geneva: World Health Organization;2019

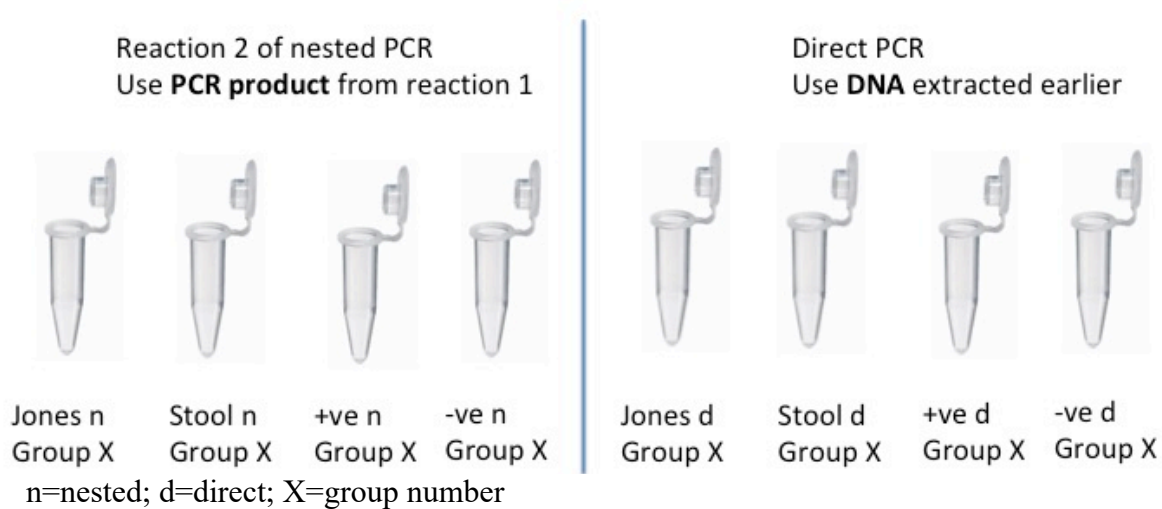
*Protocol nested PCR, reaction 2 and direct PCR reaction*

Primers used: RD5F + RDr9R

In this part of the practical, you will set up reaction 2 of the nested PCR and a direct PCR reaction. For the first one, the DNA template will be the PCR product from reaction 1 of the nested PCR. For the direct PCR, the DNA template will be the DNA you extracted earlier today.

You will need to set up eight tubes. Separate them in two groups of four (four on the left side of the rack, four on the right). The four tubes on one side will be used for reaction 2 of the nested PCR while the other four will be used for the direct PCR. The total volume of each reaction is 50  $\mu$ l.

Label your tubes as follows:



a. Master mix preparation in a 1.5 ml tube:

Forward primer BsRD5F: 9 x 2  $\mu$ l = 18  $\mu$ l  
 Reverse primer BhRDr9R: 9 x 2  $\mu$ l = 18  $\mu$ l  
 PCR BIO Taq Mix Red: 9 x 25  $\mu$ l = 225  $\mu$ l  
 Water: 9 x 19  $\mu$ l = 171  $\mu$ l

b. Mix well by pipetting up and down gently

c. Distribute 48  $\mu$ l of the master mix in each tube

d. Add 2  $\mu$ l of water in the –ve control tubes and close them.

e. Add DNA as follows:

- For nested PCR, reaction 2: Add **1  $\mu$ l** of PCR product (as shown in figure above) and close the tubes
- For direct PCR: Add **2  $\mu$ l** of DNA (as shown in figure above) and close the tubes

f. Finally, add 2  $\mu$ l of +ve control DNA in the tubes and close them

g. Insert tubes in PCR machine

h. Run the program. This will be pre-programmed for you. The PCR protocol is provided below for your reference

- 94°C for 5 minutes
  - 94°C for 1 minute
  - 50°C for 1 minute
  - 72°C for 1 minute
  - 72°C for 10 minutes
  - Hold at 12°C
- } 30x

Fragment size for 2nd PCR reaction: 600 bp

**Day 2 – 31<sup>st</sup> of August, 2023**

*Instructors: Dr. Anastasios Tsaousis, Prof. Funda Dođruman-Al, Prof. Ayman A. El-Badry, Dr. Eleni Gentekaki, Dr. Rune Stensvold, Dr. Kevin Tan*

09:00 – 09:45	Lecture - Introduction to terminologies (PCR vs qPCR, prevalence, survey)- <i>all instructors</i>
09:45 – 10:30	Practical - Set up qPCR- <i>Dr. Anastasios Tsaousis, Dr. Eleni Gentekaki, Prof. Funda Dođruman-Al</i>
10:30 – 11:00	Coffee break
11:00 – 11:30	Practical – agarose gel casting- <i>Dr. Anastasios Tsaousis, Dr. Eleni Gentekaki,</i>
11:30 – 12:30	Practical – Microscopy- <i>Prof. Funda Dođruman-Al, Prof. Ayman A. El-Badry</i>
12:30 – 14:00	Lunch break – Novogene demonstration
14:00 – 15:00	Practical – Gel electrophoresis (PCR and qPCR products)- <i>Dr. Anastasios Tsaousis, Dr. Eleni Gentekaki</i>
15:00 – 15:30	Practical – Microscopy (optional)- <i>Prof. Funda Dođruman-Al, Prof. Ayman A. El Badry</i>
15:30 – 16:00	Tea break
16:00 – 16:30	Lecture-Importance of <i>Blastocystis</i> culturing- <i>Dr. Kevin Tan</i>
16:30 – 17:30	Practical - Data visualisation and interpretation

## 1. qPCR

qPCR is a sensitive method used to survey for *Blastocystis* presence.

### Important to know before starting:

- Keep all primers and PCR reagents on ice.
- Wipe all pipettes with IMS before every PCR (also make sure you don't touch the pipette in any of the tubes as this could lead to contamination).
- For the negative control, replace 1  $\mu$ l of DNA with 1  $\mu$ l of milliQ sterile water (i.e. no DNA template in this control). For the positive control, a known positive sample will be provided.
- When you make the master mix, do not add the DNA. Add your sample DNA to each individual PCR tube/qPCR wells once the mastermix has been aliquoted.
- The volume included below includes the volume required for one sample. You will then need to calculate the total volume by considering the number of samples you have, including a positive and negative control, as well as an extra reaction mix. Let's say, for example, you have 2 samples, along with a positive and negative control, and an additional spare reaction mix, This implies that you will need to multiply the volume for one sample by the total number of reactions, which in this case would be 5.

### *Protocol qPCR*

Primers used: BL18SR2PP + BL18SPPF1

BL18SR2PP 5'-TCTTCGTTACCCGTTACTGC-3'

BL18SPPF1 5'-AGTAGTCATACGCTCGTCTCAAA-3'

You will run two reactions, one for stool and one for Jones. The total volume of each reaction is 10  $\mu$ l. Each reaction should contain:

- 5  $\mu$ l Luna Taq Universal MasterMix
- 0.5  $\mu$ l Primer\_R (BL18SR2PP)
- 0.5  $\mu$ l Primer\_F (BL18SPPF1)
- 2  $\mu$ l water
- 2  $\mu$ l DNA

a. Make a master mix in a 1.5 ml tube as follows:

Forward primer BL18SPPF1:	3 x 0.5 $\mu$ l = 1.5 $\mu$ l
Reverse primer BL18SR2PP:	3 x 0.5 $\mu$ l = 1.5 $\mu$ l
Luna Taq Universal MasterMix:	3 x 5 $\mu$ l = 15 $\mu$ l
Water:	3 x 2 $\mu$ l = 6 $\mu$ l

b. Mix well by pipetting up and down gently

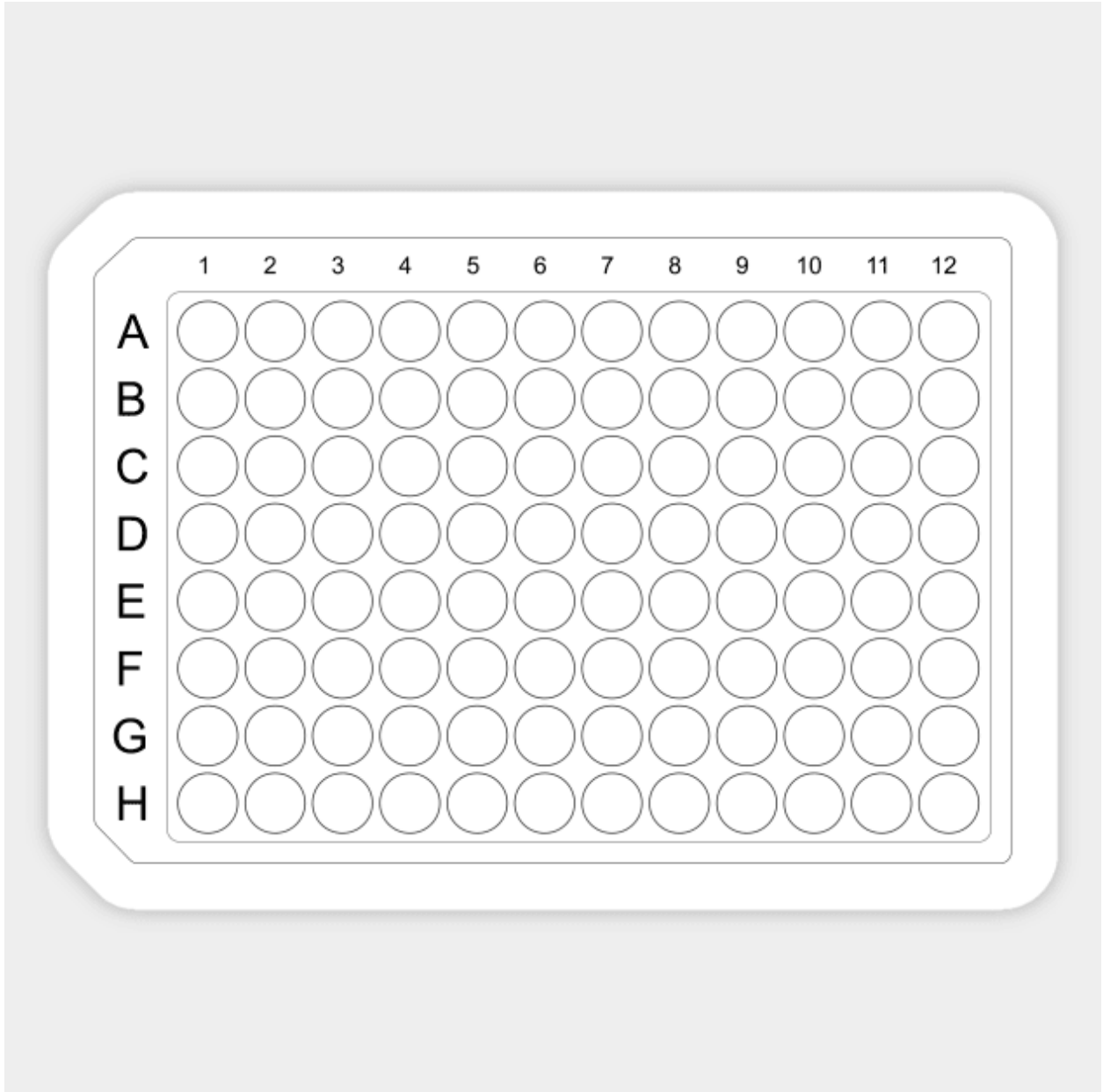
*Note: qPCR is run on a plate so samples need to be loaded in wells.*

c. Load 8  $\mu$ l of master mix per well. Since you will run two reactions (one stool, one Jones) you will use two wells.

d. Add 2  $\mu$ l of DNA in each well.

*Note: A positive and a negative control will be run for the whole plate*

e. Indicate the location and type of sample in the plate below



f. Run the program. This will be pre-programmed for you. For your reference the programme is given below:

- 95°C for 5 minutes
  - 95°C for 5 sec
  - 68°C for 10 sec
  - 72°C for 15 sec
  - 95°C for 5 sec (Dissociation or melt curve)
- } (x45)

Product size: 300 bp

## 2. Agarose gel casting

### *Protocol*

1. Prepare a 1% (w/v) agarose gel:

a) Add 0.5 g of agarose into a beaker or small flask (this has already been measured for you)

b) Add 50 ml of 1X TBE buffer

*Note: containers with TBE are located on the side benches*

c) Use the microwave to dissolve the agarose (pay close attention so it does not boil over)

d) Let it cool down a bit and add 5  $\mu$ l of dye (this has already been aliquoted for you)

e) Swirl the flask to mix the dye, and pour mixture into the gel tray, which has **two** well combs

f) Let it solidify (~30 min)

## 3. Microscopy

### *Xenic culture protocol of Blastocystis in Jones' medium*

You do not need to prepare Jones medium as this has already been prepared for you and inoculated with *Blastocystis*. The protocol for preparing the media is provided for you for future reference.

### **Reagents and preparation of the solutions**

#### **Stock solution**

Na<sub>2</sub>HPO<sub>4</sub>..... 9.46 g in 1L distilled water

KH<sub>2</sub>PO<sub>4</sub>..... 9.08 g in 1L distilled water

NaCl..... 9.00 g in 1L distilled water

#### **Preparation**

1. Mix 93.8 mL Na<sub>2</sub>HPO<sub>4</sub> with 31.3 ml KH<sub>2</sub>HPO<sub>4</sub> and 562.5 ml NaCl.

2. To the buffered solution, add yeast extract (Oxoid) to 0.1%.

3. Sterilize by autoclaving at 121°C for 15 min.

4. After cooling, store at 4°C.



5. Before use, add 10% horse serum (heat inactivated at 56°C for 30 min) to Jones' Medium.

## Procedure

1. For each faecal specimen, transfer ~50 mg of sample (formed stool) with a wooden swab to a 2 ml microcentrifuge tube containing 1.5 ml of Jones medium or pipette 0.1 ml if the sample is watery

2. The culture tubes are incubated at 37 °C for 48-72 h

3. When subculturing, transfer about 50-100 µl of the sediment from the tube used to culture into 1.5 ml fresh Jones medium containing 10% (heat-inactivated) horse serum. Subculture is made every 3 to 4 days.

*Note: The presence of bacteria in the sample will create the essential anaerobic environment for Blastocystis growth*

5. Cultures can be examined with or without iodine at x100 and x400 magnification and evaluated for the presence of especially vacuolar (the main stage seen) but also other stages of *Blastocystis*.

## References

<http://www.blastocystis.net/p/lab-stuff.html>

Jones, W. R. 1946. The experimental infection of rats with *Entamoeba histolytica*; with a method for evaluating the anti-amoebic properties of new compounds. *Annals of Tropical Medicine and Parasitology*, 40: 130-140.

## ***Protocol for observing Blastocystis in Jones media***

1. Place one drop of culture on the left side of the slide and one drop on the right side.
2. Add one drop of iodine (working solution) on the right side of the same slide.
3. Place 22x22 mm coverslips on top of the drops without creating bubbles.
4. Examine at x100 and x400 magnification

## **4. Gel electrophoresis**

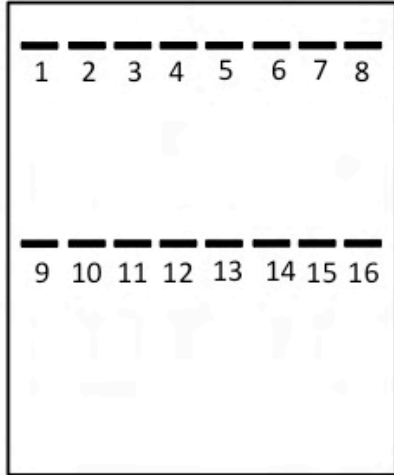
### *Protocol*

Samples to load: 4 tubes of nested PCR, **reaction 2**; 4 tubes from direct PCR; two tubes of qPCR

1. Carefully remove the combs and place the gel on the electrophoresis apparatus filled with 1X TBE buffer
2. Load samples on wells  
PCR products: load 10 µl of product per well  
qPCR products: mix the product with 2 µl of dye and load 6 µl per sample

3. Load 5 µl of ladder (tube labelled as “ladder”) in an empty well on the top and in another empty well on the bottom of the gel

Suggestion on how to load the samples



Wells 1-4: PCR products of nested PCR, reaction 2

Wells 5-6: qPCR products

Well 7: ladder

Wells 9-12: PCR products of direct PCR

Well 13: ladder

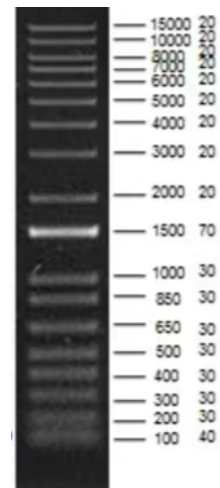
4. Run the gel at 80V for 30 min
5. Visualize and image your gel

## 5. Microscopy

During this time, you can observe the permanent trichrome stain slides. Alternatively, you could make new slides from Jones media and further observe live *Blastocystis*.

## 6. Data visualization and interpretation

Looking at and discussing gel electrophoresis and qPCR melt curves. Picture below is the ladder you used and is useful for determining the size of your bands. PCR product of nested PCR, reaction 2 size=600 bp; qPCR product=300 bp



### **Day 3 – 1<sup>st</sup> of September, 2023**

*Instructors: Dr. Rune Stensvold, Dr. Eleni Gentekaki*

*Guest lecturers: Dr. Monica Santin and Jenny Maloney*

09:00 – 10:00	Lecture - Introduction to the genetic diversity of <i>Blastocystis</i> , subtyping system, primers, host specificity - <i>Dr. Rune Stensvold</i>
10:00 – 10:30	Coffee break
11:00 – 12:00	Computer-based tutorial – Bioinformatics 1 (sequence editing and assembly) – <i>Dr. Rune Stensvold</i>
12:00 – 12:20	Lecture – Introduction to BLAST basics and MLST database - <i>Dr. Rune Stensvold</i>
12:20 – 13:00	Computer-based tutorial – Bioinformatics 2 (% identity, % cover, interpretation) – <i>Dr. Rune Stensvold</i>
13:00 – 14:00	Lunch break
14:00 – 15:30	Computer-based tutorial – Bioinformatics 3 (Genetic distance calculation; Tree construction)
15:30 – 16:00	Tea break
16:00 – 17:00	Lecture – <i>Blastocystis</i> subtyping – <i>Dr. Monica Santin and Jenny Maloney</i>
17:00 – 18:00	Discussion on <i>Blastocystis</i> subtyping
18:00	End of the day

## Bioinformatics 1

### **Assembling multiple AB1 files from the same organism (large PCR product sequenced by multiple primers)**

The STADEN software package can be downloaded free from <http://staden.sourceforge.net/> (it might be advantageous to use the Mozilla Firefox browser for the download).

Download the two AB1 files to a designated folder on your computer (name it 'OrganismX' for example).

- Open the PREGAP4 option in STADEN
- Press the "Add files" option in the menu
- Locate and upload the two AB1 files
- Press the "Configure Modules" option in the menu
- Un-tick the following:
  - Cloning Vector Clip
  - Screen for unclipped Vector
  - Sequencing Vector Clip
- Press "RUN"
- Files are now converted to ZTR and EXP files. The EXP files are used for sequence assembly. It is important that all the files (including the ABI files) are kept in the same folder, otherwise, sequence trace views will be disabled.
- Pick the 'Files to process' banner
- Press 'Clear current list'
- Press "Add files"
- Select the EXP files that you would like to upload (and assemble) (You will have to set the "Files of type" bar to .EXP files)
- Un-tick all boxes ticked by default, and tick "Gap4 Shotgun Assembly"
- Tick the "Create New Database"
- Enter a name for the Database in the designated box (Gap4 database name); e.g., 'OrganismX'
- Press "RUN"
- A database file with the suffix .0 will now be available in the folder containing the .exp files

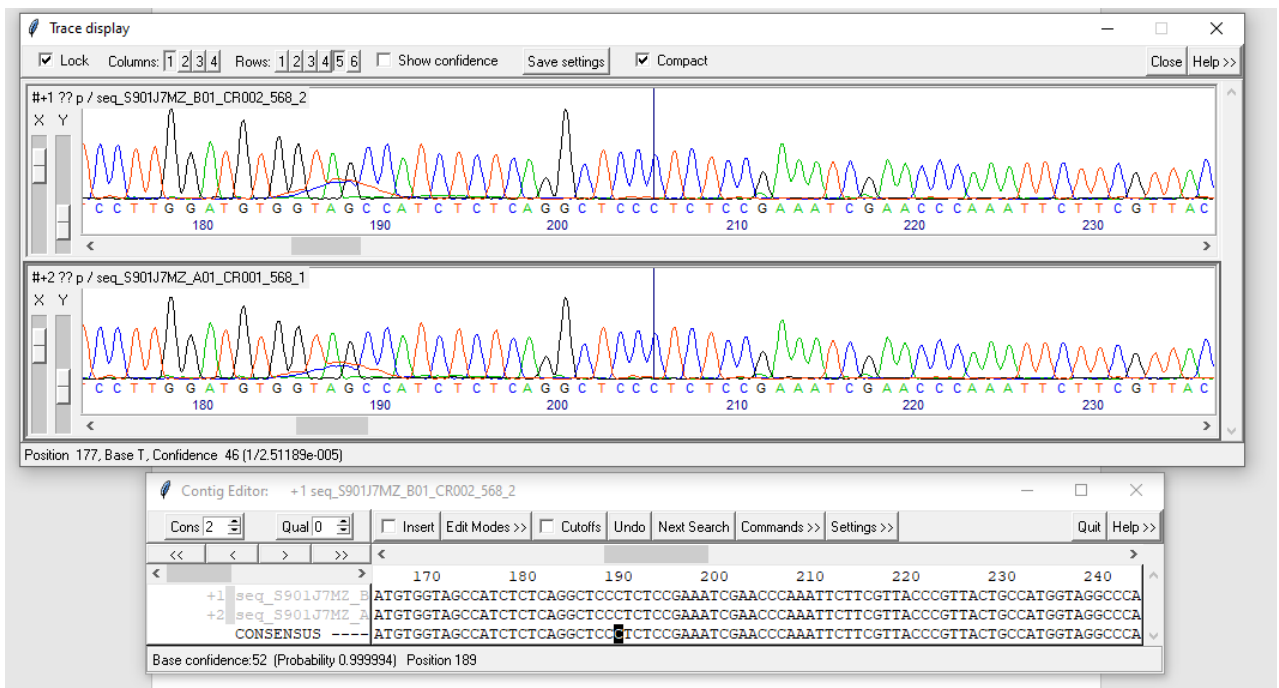
In the event that you would like to add files to an existing database (not necessary in this example, but here goes:)

- Open the database file (ending with .0)
- From the menu, choose "Assembly"
- From the drop down menu, choose "Normal Shotgun Assembly"
- Leave the default parameters, but tick "selection" instead of "file".
- Upload the .exp file in question
- In the "Save failures to" box, enter the following list name: fail
- Press "OK"

## Contig editing

Now you would want to go through and maybe edit the contig(s). Open the file with the OrganismX.0 file. You will see one contig in the contig selector in the present example. Right click on the contig, and choose “Edit contig” from the drop down menu. You can choose to look at all sequences in the contig at the same time simply by double-clicking in the consensus line. You can use the scroll bars (‘X’ and ‘Y’) to the left to optimize your view.

Getting to know the program takes a little practice. Also, the sequences in this example are of sub-optimal quality (note the presence of ‘ethanol blobs’), so it will take some time to work out the bases, and under normal circumstances, only high-quality sequences should be used. But the use of poor-quality sequences can be good for teaching, since it makes you explore some of the editing options available in the software.



The screenshot shows two windows from the Blastocystis software. The top window, titled "Trace display", shows two chromatograms (labeled #1 and #2) with a consensus sequence below them. The consensus sequence is: C C T T G G A T G T G G T A G C C A T C T C T C A G G C T C C C T C T C C G A A A T C G A A C C C A A A T T C T T C G T T A C. The bottom window, titled "Contig Editor", shows a list of sequences and a consensus line. The consensus line is: ATGTGGTAGCCATCTCTCAGGCTCCCTCCGAAATCGAACC CAAATTCCTCGTTACCCGTTACTGCCATGGTAGGCCCA. The cursor is positioned at the 189th base of the consensus line.

## A few tips on sequence editing

When you want to edit your sequences, make sure you are in “**Edit mode 2**” by simply selecting it from the drop-down menu in “Edit Modes”. In this menu, you will have to actively tick “Allow del any in cons” if you want to delete bases.

A very valuable feature is the “**Cutoffs**” feature, which enables you to see the “hidden” sequence bases (considered under the quality threshold). By ticking the “Cutoffs” box, these “hidden” sequence parts will become visible. It is possible to integrate them as “valid” sequence parts simply by putting the cursor at a given spot, and then press the left or the right arrow (< or > on the keyboard) all depending on whether you are at the 5’ end of the sequence (<) or the 3’ end (>).

Remember that you can always ‘Undo’ if you make a change that is not intended. Also, if you make intended changes, remember to save them (you can do this at the end, by clicking the red cross, then a dialogue box will appear asking you whether you want to save changes or not).

In the present example there are three contigs in the beginning, - eventually, you want to be able to join these contigs as best as possible to produce one (almost) complete SSU rDNA sequence. The short contig can be edited to include many more bases, and once this has been taken care of, the contigs can be joined in order to produce one full sequence (reflecting the entire SSU rRNA gene). As mentioned, the short contig is tricky and requires a lot of work due to poor sequence quality. You may want to try and use the reference sequences for support (look for conserved areas in MEGA sequence alignment, if you are ‘missing’ some bases, and then enter them in the contig where relevant).

To check for potential **contig joins**, go to

- “View” in the main menu
- Select the “Find internal joins” in the drop down menu
- Use the default parameters and press “OK”
- Any potential joins will now appear as diagonal bars on the “Contig Comparator” board.
- Click on a bar and to ensure proper alignment, un-tick the “lock” box, tick the “align” box and then re-tick the “lock” box.
- Any discrepancies (mismatches, including gaps) will be marked by exclamation marks.
- Attend to these discrepancies and make sure you are in edit mode 2 (in both contigs)
- You can insert bases by clicking the “Insert” box (remember to un-tick it immediately afterwards unless you have a series of insertions).
- NOTE: If you insert a base, sometimes it is useful to “hide” the sequence by the Shift and Arrows keyboard buttons, otherwise the whole sequence will be shifted a base, which may have a huge influence on the rest of the contig.
- You can look for sequence mismatches within a contig by selecting the Settings button in the main menu, and choose from the “Highlight Disagreements” menu. This enables a fast check on potential mismatch problems.
- You can look for particular sequence “stretches” (e.g. primer annealing sites) in a given contig by selecting “Next search” in the main menu in an open contig and then tick the “sequence” box.

If in a contig you have one sequence that is “out of sync” you can remove it from the contig into a separate contig (in the same database) by

- left-clicking on the name of the sequence in the panel to the left in the contig window.
- Then right click and select “Remove reading (this only).”
- Close the window and you will be asked to choose from a variety of options, I tend just to go with the default option (move reading(s) to new contig(s))
- Then if you want, you can always add the sequence to the contig by looking for potential joins, and then if you make the join, the sequence will (probably) be incorporated (correctly) into the contig from which you just removed it.

### **Copying a sequence to clip board**

If you want to copy a sequence in a contig to the clip board (for instance if you need to do a BLAST search), you simply enter the cursor at one of the ends of the sequence, left click on it

and keeping the left click button pressed, you drag it towards the end of the sequence of interest and simply paste it into the box at

[http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST\\_PROGRAMS=megaBlast&PAGE\\_TYPE=BlastSearch&SHOW\\_DEFAULTS=on&LINK\\_LOC=blasthome](http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST_PROGRAMS=megaBlast&PAGE_TYPE=BlastSearch&SHOW_DEFAULTS=on&LINK_LOC=blasthome) as usual.

## Bioinformatics 2

### EXAMPLES FOR BLASTing EXERCISES

1. Try BLASTing the following sequence against

1) NCBI Database

[https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST\\_SPEC=GeoBlast&PAGE\\_TYPE=BlastSearch](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST_SPEC=GeoBlast&PAGE_TYPE=BlastSearch)

>BlastSeq1

```
CTGCCAGTAGTCATACGCTCGTCTCAAAGATTAAGCCATGCATGTGTAAGTCTAAATAGTTTTA
GTGTGAAACTGCGAATGGCTCATTATATCAGTTATAGTTTATTTGGATAAAAGTTTCTATGGATA
ACCGTAGTAATTCTAGGGCTAATACATGTCAAATCTCTTTGAGATTTCTTATTAGAATTAAC
CAATGTTATTAGTGAGTAATAATAATGTAATCTATCGCATGCTTTAGTAGCGATAATCCTTTCAA
GTTTCTGCCCTATCAGCTTTCGATGGTAGTGTATTGGACTACCATGGCAGTAACGGGTAACGAA
GAATTTGGGTTTCGATTTTCGGAGAGGGAGCCTGAGAGATGGCTACCACATCCAAGGAAGGCAG
CAGGCGCGCAAATTACCCAATCCTGACACAGGGAGGTAGTGACAATAAATCACAATGCGGAA
CCTTTGGTTTTGCAATTGGATTGAGAAGAAGGTACAAATGATATCGATAAACAATTGGAGGGC
AAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTATATTAACGTTGTTGCAGTT
AAAAAGCTCGTAGTTGAAGTGAGATGAGTATTTTATAAGAATGTCAAATCTTGTA
```

Comment on your findings:

Can you identify the subtype? How certain are you? Reason your statement.

2. Try BLASTing the following sequence against the NCBI Database

>BlastSeq2

```
GATTTATCTGGTTGATCCTGCCAGTAGTCATACGCTCGTCTCAAAGATTAAGCCATGCATGT
GTAAGTCTAAATAGTTTTAGTGTGAAACTGCGAATGGCTCATTATATCAGTTATAGTTTATTT
GGATGAATTTTTCTATGGATAACCGTAGTAATTCTAGGGCTAATACATGTGTGAATCTCTTTG
AGATTTTCTTATTAGAATAAAAACCAATTGTATTAGTGAGTAATAATAATAGAATCGATCGCA
TGCTTTTGTAGCGATAATCCTTTCAAGTTTCTGCCCTATCAGCTTTCGATGGTAGTGTATTGG
ACTACCATGGCAGTAACGGGTAACGAAGAATTTGGGTTTCGATTTTCGGAGAGGGAGCCTGAG
AGATGGCTACCACATCCAAGGAAGGCAGCAGGCGCGCAAATTACCCAATCCTGACACAGGG
AGGTAGTGACAATAAATCACAATGCGGAACCTTTGGTTTTGCAATTGGATTGAGAAGAAGG
TACAAATCATATCGATAAACAATTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCC
AGCTCCAATAGCGTATATTAACGTTGTTGCAGTTAAAAAGCTCGTAGTTG
```

Comment on your findings:

Can you identify the subtype? How certain are you? Reason your statement.

What could you do to add more support for calling the subtype?



3. Try BLASTing the following sequence against the NCBI Database

>BlastSeq3

```
ACAATGCGGGACTTTTAGTTTTGCAATTGGATTGAGAACAATGTAAAAGAGTTATCG
ATAAACAATTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAAT
AGCGTATATTAACGTTGTTGCAGTTAAAAAGCTCGTAGTTGAAGTGTGGGGTGAAA
AGTTGGTTGTTTCTTTGAAATGGCTGTCTGCTCACCCCCTCTCCAATGTGTTATTATG
GGTATTAATTTACTTGTAGTAGCTATTGGTCATTTACTGTGAGAAAATTAGAGTGTTT
AAAGCAGGCATTTGCTTGAATATATTAGCATGGAATAATAGTAAAAGGCTTTTGTGG
TCGATTTAATTGGTTTGTTTCATAAAAGCAAGATTAATAGGGACAGTTGGGGGTATTC
ATATTCAATAGTTAGAGGTGAAATTCTCGGATTTATGGAA
```

Can you identify the subtype? How certain are you?

What could you do to add more support for calling the subtype?

4. Try BLASTing the whole following sequence against the NCBI Database.

>BlastSeq4

```
AGATTAGCCATGCATGTGTAAGTATAAATAATTTTACTTTGAAACTGCGAATGGCTCATTAT
ATCAGCTATAGTTTATTAGAAATACTTTACTACTTGGATAACCGTAGTAATTCTAGGGCTAAT
ACATGAATAAATCCTCAACTTCGGAAGAGGTGTGTTTATTACTATGAAAACCATATATCTTC
GGATATAATAGTGAGTAATAAACAACCTTCGTATCGTATGTTCGTACGGATAACGATAGTTCT
TTCAAGTTTCTGCCCTATCAGCTTCGATGGTAGTATATGGGACTACCATGGCAGTAACGGG
TAACGAAGAATTTGGGTTTCGATTCGGAGAGGGAGCCTGAGAGATGGCTACCACATCCCAG
GAAGGCAGCAGGCGCGTAAATTACCCAATCCTGACATAGGGAGGTTAGT
```

GACAATAAATCAC  
ATGCGGGACTTCAGTCTTGCAATTGGATTGAGAACAATGTACAACCTTATCGATAAGCCAT  
TGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATCCAGCTCCAATAGCGTATATTAACGT  
TGTTGCAGTTAAAAAGCTCGTAGTTGAAGTGTGGGTGATCGCTGTTGTGAGACTTCGGTCTC  
TCGACAGTAAGTCACCCCCTTCCAGTATCCAGTAGTGGGTATTCAGTTACTTACTATTGTGTG  
TTGGTCGTTTACTGTGAGAAAATTAGAGTGTTCAAAGCAGGCGTTTGCTTGAATAGATTAGC  
ATGGAATAATAATTGAAGGCTTTTCGTGTTTCCATATGATTGGTTTGTTCATAAAAGCAAGAT  
TAAAAGGAACAGTTGGGGGTATTCATATTCAATAGTTAGAGGTGAAATTCTCGGATTTATGG  
AAGATGAACTAGTGCGAAAGCATTTACCAAGGATGTTTTTCATTAATCAAGAACGAAAGCTA  
GGGGATCGAAGAGGCTTAGATACCCTCGTAGTCTTAGCTATAAACTATAACCGACTAGGGGTT  
AGTAGAAGTCATTGTGTTTCTATTAGTACCTTATGAGAAATCAAAGTCTTTGGGTTCCGGGG  
GGAGTATGGTCGCAAGGCTGAAACTTAAAGGAATTGACGGAAGGGCACCACCAGGAGTGG  
AGCCTGCGGCTTAATTTGACTCAACACGGGAAACTTACCAGGTCCAGACATAGGAAGGATT  
GACAGATTGATAGCTCTTTCTTGATTCTATGGGTGGTGGTGCATGGCCGTTCTTAGTTGGTGG  
AGTGATTTGTCTGGCTAATTCCGATAACGAACGAGACCTCCGCCTTAACTAGTGACGTGTA  
TTGTGATATGCGTTGCTTCTTATAGGGACACTATATGTAAAATGTAGGGAAGCTGGAGGCAA  
TAACAGGTCTGTGATGCCCTTAGATGTTCTGGGCTGCACGCGCGGACACTGATCTATTCAA  
CGAGTGGCTGGGTCGAGAGACTTGGCAAATCTTGTAAGTAGATCGTGATGGGGATTGAGG  
CTTGAAATTTTTCATCATGAACGAGGAATTCCTAGTAAACGCAAGTCATCAACTTGCATTGA  
TTACGTCCCTGCCCTTGTACACACCGCCCGTCGCACCTACCGATTGAATGATCCGGTGAAC  
ACTTTGGATTTTCTATTTTTTAATTGTAACAGATTAAGAATAGAAGAGAAGTCGTGTAAAT  
CTTATCATTTAGAGGAAGGTGAAGTCGTAACAAGGT

Comment on your finding! Which subtype is it?

Try BLASTing the individual, coloured parts of the sequences one by one using NCBI and PubMLST databases. Comment on your findings.

What do you make of this DNA sequence entry?

5. Try BLASTing the following sequence against the NCBI Database.

>BlastSeq5

```
ACCTGGTTGATCCTGCCAGTAGTCATACGCTCGTCTCAAAGATTAAGCCATGCATGTGTAAG
TATAAATAATTCTACTTTGAAACTGCGAATGGCTCATTATATCAGTTATAGTTTATTAGAAAA
AATGTACTACTTGGATAACCGTAGTAATTCTAGGGCTAATACATGAGTAAATCCTCGACTTT
GAAGAGGTGTGTTTATTAGTATGAAAACCATATGCTTCGGCATAAATAGTGAGTAATAATAAC
CTTTCATATCGTATGTTGAAATAACGATAGTTCTTTCAAGTTTCTGCCCTATCAGCTTTCGAT
GGTAGTGCATTGGACTACCATGGCAGTAACGGGTAACGTTGACGGAAGGGCACCACCAGGA
GTGGAGCCTGCGGCTTAATTTGACTCAACACGGGGAAACTTACCAGGTCCAGACATAGGAA
GGATTGACAGATTGATAGCTCTTCTTGATTCTATGGGTGGTGGTGCATGGCCGTTCTTAGTT
GGTGGAGTGATTTGTCTGGTTAATTCCGATAACGAACGAGACCTCCGCCTGCTAACTAGGTT
TTATTAATTTTAGTAAAATCTTCTTAGAGGGACACTATACGTTTTAGTGTAGGGAAGCTGG
AGGCAATAACAGGTCTGTGATGCCCTTAGATGTTCTGGGCTGCACGCGCGGACACTGATCT
ATCAACGAGTGACTGGATCGATAGATTTGGTAAATCTTTTGAAAATAGATCGTGCTGGGGA
TTGATGCTTGAAATTTTTCATCATGAACGAGGAATTCCTAGTAAACGCAAGTCATCAACTTG
CATTGATTACGTCCCTGCCCTTGTACACACCGCCCGTCGCACCTACCGATTGAATGATCCGG
TGAACACTTTGGATTTTGGCTAGTTGACTGTAACAGGTCATTTAGCATAAGAGAAGTCGTGT
AAATCTTATCATTTAGAGGAAGGTGAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCAGA
AGGATC
```

Comment on your findings

Useful Links

- Blastocystis full-length SSU rDNA reference sequences  
<http://entamoeba.lshtm.ac.uk/ref.blasto.txt>
- Staden Package <https://staden.sourceforge.net/> - look up recent releases and pick the link that matches your requirements/specifications
- Nucleotide BLAST  
[https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST\\_SPEC=GeoBlast&PAGE\\_TYPE=BlastSearch](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST_SPEC=GeoBlast&PAGE_TYPE=BlastSearch)
- MEGA7 <https://www.megasoftware.net/>
- Blastocystis PubMLST database <https://pubmlst.org/organisms/blastocystis-spp>
- YouTube tutorial to PubMLST database:  
<https://www.youtube.com/watch?v=Isj7zgBJzDI&t=1s>
- MultAlin <http://multalin.toulouse.inra.fr/multalin/>
- Clustal Omega <https://www.ebi.ac.uk/Tools/msa/clustalo/> (can be used with JalView)
- Jalview <https://www.jalview.org/>

Bioinformatics 3

## Genetic distance calculation

This calculation is useful for determining the %genetic distance (i.e. how many base pairs are different) between sequences. Multiple sequences can be compared at the same time. MEGA is used to perform the calculation.

1. Open MEGA7.0
2. Open the ts2\_gen\_dist.fasta file using a text editor and copy all sequences to the clipboard
3. Choose the “Align” option on the top left of menu bar in MEGA7.0 and choose the “Edit/Build Alignment” option
4. Select “Create a new alignment” option then press OK
5. Select “DNA” when asked: are you building a DNA or protein sequence alignment?
6. A large window will appear. Click “edit” on the top menu bar and select “Paste”. All sequences will appear
7. Align all sequences by pressing the muscle arm button (located beside the W button in the menu bar), select “Align DNA” and then “Compute”
8. Look at the beginning (5’) and the end (3’) of the alignment and notice that some sequences are longer than others. For this type of analysis, gaps at the beginning and at the end of the alignment must be deleted, gaps within are ok.

Species/Abbrv	Group Name	T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
U51151_Blastocystis_ST1_ref		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
EU445488_Blastocystis_M2_ST1		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
EU445492_Blastocystis_H4_ST1		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
AB107987_Blastocystis_MJ99_424_ST1		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
EU445490_Blastocystis_M9_ST1		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
EU445486_Blastocystis_P3_ST1		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
AB091234_HJ96A-26_cloneb_ST3_ref		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
JX132219_CEss532_ST3_ref		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
AB107965_CJ99-363_ST3_ref		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
KX618192_ST3_ref		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
AY244620_H900-12_ST4_ref		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
GU992412_JEB-2010_isolate_I6_ST2_ref		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
AB107962_Blastocystis_PJ99_172_ST1_ref		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
AY618266_Blastocystis_989_ref		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
AB070997_Blastocystis_MJ92_2_ST2_ref		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
AB070987_Blastocystis_ST2_ref		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
AB107969_Blastocystis_MJ99_116_ST2		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
EU427518_Blastocystis_MWJ03_1_ST2		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....
AY618265_Blastocystis_84_ST2		T	A	A	T	C	T	T	A	T	C	A	T	T	T	A	G	A	G	G	G	T	G	A	A	G	T	C	G	T	A	A	C	A	A	G	G	T	T	C	C	G	T	A	G	G	T	G	A	A	C	C	T	G	C	G	G	A	A	G	G	A	T	C	A	.....

In the picture above, positions within the black square must be deleted so that there are no gaps in the end of the alignment. Do the same for the beginning of the alignment

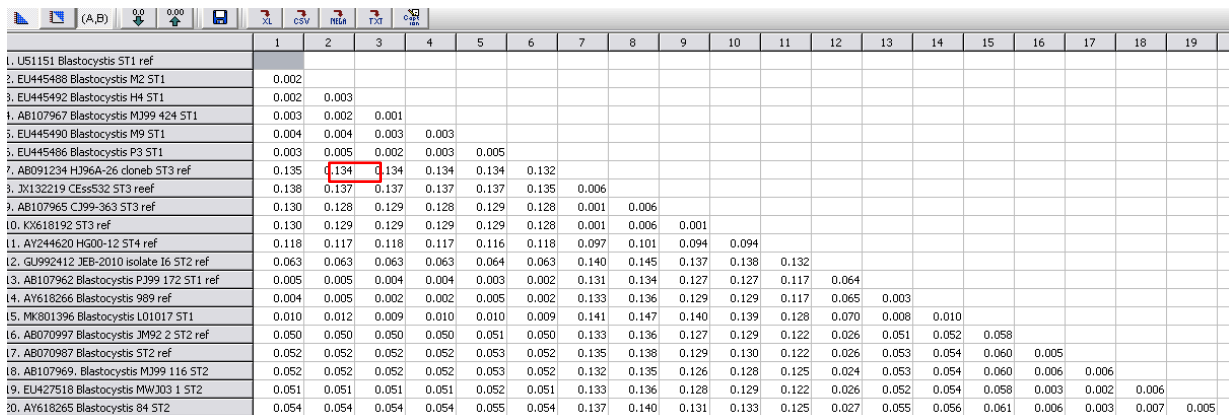
9. Select “Data” on the top left of the MEGA7.0 menu, then “Export alignment”, then MEGA format
10. A window will appear asking you to name and save your file.
11. Select “No” to the question: “protein-coding nucleotide sequence data?”
12. Select “Data” on the top left of the MEGA7.0 menu, then “Exit alignment explorer”
13. Next, perform the genetic distance analysis by selecting “Distance”

14. A window will appear asking you to select the appropriate file (in meg format, created in step 9), then press “Open”

15. Another window will open. Select “Kimura 2-parameter model” under Model/Method, “Gamma Distributed (G)”, under Rates among Sites and “Pairwise deletion” under Gaps/Missing Data Treatment”, then press “Compute”

*Note: feel free to change the parameters (e.g. the Gaps/Missing data treatment options) and see how this changes the calculated values*

16. A table will appear.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
1. U51151 Blastocystis ST1 ref																				
2. EU445488 Blastocystis M2 ST1	0.002																			
3. EU445492 Blastocystis H4 ST1	0.002	0.003																		
4. AB107967 Blastocystis M399 424 ST1	0.003	0.002	0.001																	
5. EU445490 Blastocystis M9 ST1	0.004	0.004	0.003	0.003																
6. EU445486 Blastocystis P3 ST1	0.003	0.005	0.002	0.003	0.005															
7. AB091234 H396A-26 cloneb ST3 ref	0.135	0.134	0.134	0.134	0.134	0.132														
8. JX132219 CEse532 ST3 reef	0.138	0.137	0.137	0.137	0.137	0.135	0.006													
9. AB107965 C399-363 ST3 ref	0.130	0.128	0.129	0.128	0.129	0.128	0.001	0.006												
10. KY618192 ST3 ref	0.130	0.129	0.129	0.129	0.129	0.128	0.001	0.006	0.001											
11. AY244620 HG00-12 ST4 ref	0.118	0.117	0.118	0.117	0.116	0.118	0.097	0.101	0.094	0.094										
12. GU992412 JEB-2010 isolate I6 ST2 ref	0.063	0.063	0.063	0.063	0.064	0.063	0.140	0.145	0.137	0.138	0.132									
13. AB107962 Blastocystis P399 172 ST1 ref	0.005	0.005	0.004	0.004	0.003	0.002	0.131	0.134	0.127	0.127	0.117	0.064								
14. AY618266 Blastocystis 989 ref	0.004	0.005	0.002	0.002	0.005	0.002	0.133	0.136	0.129	0.129	0.117	0.065	0.003							
15. MK801396 Blastocystis L01017 ST1	0.010	0.012	0.009	0.010	0.010	0.009	0.141	0.147	0.140	0.139	0.128	0.070	0.008	0.010						
16. AB070997 Blastocystis JM92 2 ST2 ref	0.050	0.050	0.050	0.050	0.051	0.050	0.133	0.136	0.127	0.129	0.122	0.026	0.051	0.052	0.058					
17. AB070987 Blastocystis ST2 ref	0.052	0.052	0.052	0.052	0.053	0.052	0.135	0.138	0.129	0.130	0.122	0.026	0.053	0.054	0.060	0.005				
18. AB107969 Blastocystis M399 116 ST2	0.052	0.052	0.052	0.052	0.053	0.052	0.132	0.135	0.126	0.128	0.125	0.024	0.053	0.054	0.060	0.006	0.006			
19. EU427518 Blastocystis MWJ03 1 ST2	0.051	0.051	0.051	0.051	0.052	0.051	0.133	0.136	0.128	0.129	0.122	0.026	0.052	0.054	0.058	0.003	0.002	0.006		
20. AY618265 Blastocystis 84 ST2	0.054	0.054	0.054	0.054	0.055	0.054	0.137	0.140	0.131	0.133	0.125	0.027	0.055	0.056	0.061	0.006	0.003	0.007	0.005	

Let’s see how this table is read. We will focus on the value in the red box. Notice that each sequence has a number beside it. Sequence number 6 (EU445486) is 0.2% different than sequence number 3 (EU445492)

17. You can export this table as is by pressing the disc icon then selecting “XL: Microsoft excel workbook” under Output Format and “Matrix” under Export type. Alternatively, you can select “Column” instead of Matrix. Can you see the differences?

## Phylogeny

Next you will generate a neighbor joining (NJ) tree based on distance analysis using the same model as the one in the previous exercise. The difference here is the output, which is a tree.

1. Open MEGA7.0
2. Open the ts2\_phylogeny.fasta file using a text editor and copy all sequences to the clipboard
3. Choose the “Align” option on the top left of menu bar in MEGA7.0 and choose the “Edit/Build Alignment” option
4. Select “Create a new alignment” option then press OK
5. Select “DNA” when asked: are you building a DNA or protein sequence alignment?
6. A large window will appear. Click “edit” on the top menu bar and select “Paste”. All sequences will appear
7. Align all sequences by pressing the muscle arm button (located beside the W button in the menu bar), selecting “Align DNA” and then “Compute”

8. Select “Data” on the top left of the MEGA7.0 menu, then “Export alignment”, then MEGA format
9. A window will appear asking you to name and save your file. Suggested file name: ts2\_phylogeny\_ali.meg
10. Select “no” to the question: “protein-coding nucleotide sequence data?”
11. Select “Data” on the top left of the MEGA7.0 menu, then “Exit alignment explorer”
12. Next, build the NJ tree by selecting “Phylogeny”, then “Construct/Test Neighbor-Joining tree”
13. Open the MEGA format alignment from step 9 (ts2\_phylogeny\_ali.meg)
14. Select “Bootstrap method” under Test of phylogeny, “Kimura 2-parameter model” under Model/Method, “Gamma Distributed (G)”, under Rates among Sites and “Pairwise deletion” under Gaps/Missing Data Treatment”, then press “Compute”. Save tree as NJtree.tre
15. Proceed to maximum likelihood tree building using IQtree on the following website: <http://iqtree.cibiv.univie.ac.at/>
16. Open the ts2\_phylogeny.fasta file using a text editor and copy all sequences to the clipboard
17. Choose the “Align” option on the top left of menu bar in MEGA7.0 and choose the “Edit/Build Alignment” option
18. Select “Create a new alignment” option then press OK
19. Select “DNA” when asked: are you building a DNA or protein sequence alignment?
20. A large window will appear. Click “edit” on the top menu bar and select “Paste”. All sequences will appear
21. Align all sequences by pressing the muscle arm button (located beside the W button in the menu bar), selecting “Align DNA” and then “Compute”
22. Select “Data” on the top left of the MEGA7.0 menu, then “Export alignment”, then fasta format (suggested name: ts2\_phylogeny\_IQ.fasta)
23. Upload your alignment file in Fasta format
24. Select sequence type as DNA or just leave the default (auto-detect)
25. Substitution model should be left at “auto”, bootstrap analysis set at “ultrafast” with maximum iterations set at “1000”
26. Select “No” for SH-aLRT branch test
27. Press “Submit job”. Providing email is optional
28. Tree files can be visualized in Figtree.