

## Instructions for Obtaining a Seed Sequence for Organelle Assembly

We will be using [NOVOPlasty](#) ([Dierckxsens et al., 2017](#)) to assemble organelle genomes for our species! As we talked about at the end of Week 8 Lab, NOVOPlasty is a seed-extend based assembler. This means that the software starts with a “seed” sequence and attempts to find overlapping reads that can extend the sequence in both directions. The seed is just a string of DNA nucleotides that belong to the organelle genome of a species that is closely related to your sample. The seed sequence does not become part of the assembly, it is just used as a starting point to retrieve the first sequencing read from the raw sequencing read data set. Remember that our sequencing read files will include reads from both nuclear and organellar genomes. We don’t want to include reads from the nuclear genome in our organelle assemblies!

The protocol for obtaining seed sequences will differ slightly depending on whether your species is a plant or an animal. For animals, you will only be assembling the mitochondria genome. For plants, you must assemble the chloroplast genome before you can assemble the mitochondrial genome.

### Animals

Look for the sequence of a conserved mitochondrial gene in a species closely related to your species! A good place to start would be the cytochrome c oxidase subunit I (COI) gene. See the instructions below to learn how to do this.

### Plants

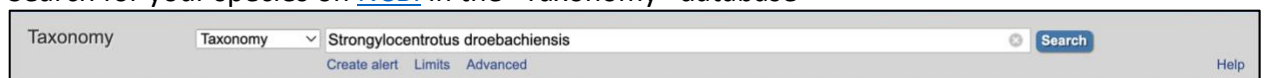
Look for a chloroplast-specific gene in a closely related species. In plants, the chloroplast and mitochondria have many similar sequences due to intergenomic transfer. It is important to find a sequence that is only found in the chloroplast, and not in the mitochondria. A good place to start would be the [RuBisCO](#) enzyme. Data on the RuBisCO enzyme for your species might be under several different synonymous names:

- Ribulose-1,5-bisphosphate carboxylase-oxygenase
- Rubisco
- RuBPCase
- RuBPco

### How to Find a Seed Sequence

I will use *Strongylocentrotus droebachiensis*, the green sea urchin, to demonstrate how to follow this protocol.

1. Search for your species on [NCBI](#) in the “Taxonomy” database



The screenshot shows the NCBI Taxonomy search interface. It features a search bar with the text "Strongylocentrotus droebachiensis" entered. To the left of the search bar is a dropdown menu labeled "Taxonomy". To the right of the search bar is a blue "Search" button. Below the search bar are links for "Create alert", "Limits", and "Advanced". In the bottom right corner, there is a "Help" link.

## 2. Click on your species

Display Settings: ▾ Summary Send to: ▾

[Strongylocentrotus droebachiensis](#)  
(green sea urchin), species, sea urchins  
[Nucleotide](#) [Protein](#)

3. Look at the info under “Entrez records”. Entrez is a database for molecular sequence data. Is there an assembly available (a row titled “Assembly” with a number greater than zero next to it)? If so, skip to step 4. In the example below, there is not an assembly available for *Strongylocentrotus droebachiensis*.

Look at the “Lineage” information. We are going to go up the lineage of your organism to hopefully find the most closely related species with an assembled genome. Click the last link in the lineage list (this most likely corresponds to a genus).

**Strongylocentrotus droebachiensis**  
Taxonomy ID: 7671 (for references in articles please use NCBI:txid7671)  
current name  
*Strongylocentrotus droebachiensis* (Muller, 1776)  
Genbank common name: green sea urchin  
NCBI BLAST name: sea urchins  
Rank: species  
Genetic code: [Translation table 1 \(Standard\)](#)  
Mitochondrial genetic code: [Translation table 2 \(Echinoderm Mitochondrial; Flatworm Mitochondrial\)](#)  
Other names:  
heterotypic synonym  
*Strongylocentrotus droebachiensis*

**Entrez records**

Database name	Direct links
Nucleotide	509
Protein	520
Genome	1
Popset	23
PubMed Central	281
Gene	41
SRA Experiments	44
Protein Clusters	13
Identical Protein Groups	193
Bio Project	2
Bio Sample	44
Taxonomy	1

Lineage (full)  
cellular organisms; Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Deuterostomia; Echinodermata; Eleutherozoa; Echinozoa; Echinoidea; Euechinoidea; Echinacea; Camarodonta; Echinidea; Strongylocentrotidae; Strongylocentrotus

After clicking the link, you will see all of the organisms that fall under that lineage. Click on the genus name one more time. In this case, it is *Strongylocentrotus*.

Lineage (full): cellular organisms; Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Deuterostomia; Echinodermata; Eleutherozoa; Echinozoa; Echinoidea; Euechinoidea; Echinacea; Camarodonta; Echinidea; Strongylocentrotidae

- Strongylocentrotus Click on organism name to get more information.
  - Strongylocentrotus droebachiensis (green sea urchin)
  - Strongylocentrotus intermedius
  - Strongylocentrotus pallidus (pale sea urchin)
  - Strongylocentrotus polyacanthus
  - Strongylocentrotus purpuratus (purple sea urchin)
- unclassified Strongylocentrotus
  - Strongylocentrotus sp. 17\_ECHINO\_001\_069
  - Strongylocentrotus sp. 17\_FA\_006
  - Strongylocentrotus sp. EAC01
  - Strongylocentrotus sp. TC-2019

- Look at the info under “Entrez” records. Is there an assembly available? If the answer is no, repeat steps 3 and 4 until you get to a taxonomic group that has an assembly available.

In this case, the Entrez records indicate that there is at least one assembly available!  
Click on the underlined number next to “Assembly.”

**Strongylocentrotus**

Taxonomy ID: 7664 (for references in articles please use NCBI:txid7664)

current name

**Strongylocentrotus**

NCBI BLAST name: **sea urchins**

Rank: **genus**

Genetic code: [Translation table 1 \(Standard\)](#)

Mitochondrial genetic code: [Translation table 2 \(Echinoderm Mitochondrial; Flatworm Mitochondrial\)](#)

Lineage (full)

cellular organisms; Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Deuterostomia; Echinodermata; Eleutherozoa; Echinozoa; Echinoidea; Euechinoidea; Echinacea; Camarodonta; Echinidea; Strongylocentrotidae

**Entrez records**

Database name	Subtree links	Direct links
Nucleotide	<a href="#">304,290</a>	-
Protein	<a href="#">42,296</a>	-
Structure	<a href="#">6</a>	-
Genome	<a href="#">4</a>	-
Popset	<a href="#">82</a>	<a href="#">32</a>
Conserved Domains	<a href="#">1</a>	-
GEO Datasets	<a href="#">679</a>	-
PubMed Central	<a href="#">3,666</a>	<a href="#">3,534</a>
Gene	<a href="#">66,362</a>	-
SRA Experiments	<a href="#">3,332</a>	-
Protein Clusters	<a href="#">13</a>	-
Identical Protein Groups	<a href="#">63,460</a>	-
Bio Project	<a href="#">94</a>	<a href="#">1</a>
Bio Sample	<a href="#">909</a>	-
Bio Systems	<a href="#">269</a>	-
Assembly	<a href="#">6</a>	-
Probe	<a href="#">99</a>	-
PubChem BioAssay	<a href="#">27</a>	-
Taxonomy	<a href="#">11</a>	<a href="#">1</a>

External Information Resources (NCBI LinkOut)

- A list of published genome assemblies will appear. Click on one of the listed assemblies. If many options appear, note that you can filter using options on the left side bar.

Organism group

Animals (1)

Customize ...

Summary

Filters activated: Latest, Exclude anomalous. [Clear all](#) to show 6 items.

[Spur\\_5.0](#)

Organism: Strongylocentrotus purpuratus (purple sea urchin)

Sex: male

Submitter: Baylor College of Medicine

Date: 2019/09/06

Assembly level: Scaffold

Genome representation: full

RefSeq category: representative genome

GenBank assembly accession: GCA\_000002235.4 (**latest**)

RefSeq assembly accession: GCF\_000002235.5 (**latest**)

IDs: 4536931 [UID] 13654338 [GenBank] 13759658 [RefSeq]

Send to:

**Assembly Information**

Assembly Help

Assembly Basics

NCBI Assembly Data Model

**Related Information**

BioProject

BioSample

Genome

Nucleotide INSDC

Shown below is the information for the latest genome assembly of the purple sea urchin (*Strongylocentrotus purpuratus*), Spur\_5.0.

Full Report ▾
Send to: ▾

### Spur\_5.0

**Organism name:** [Strongylocentrotus purpuratus \(purple sea urchin\)](#)

**Isolate:** Spur 01

**Sex:** male

**BioSample:** [SAMN00829422](#)

**BioProject:** [PRJNA10736](#)

**Submitter:** Baylor College of Medicine

**Date:** 2019/09/06

**Assembly level:** Scaffold

**Genome representation:** full

**RefSeq category:** representative genome

**GenBank assembly accession:** GCA\_000002235.4 (latest)

**RefSeq assembly accession:** GCF\_000002235.5 (latest)

**RefSeq assembly and GenBank assembly identical:** no ([hide details](#))

- Only in RefSeq: chromosome MT (in non-nuclear assembly-unit)
- Data displayed for RefSeq version

**WGS Project:** [AAGJ06](#)

**Assembly method:** Canu v. 1.8

**Expected final version:** yes

**Genome coverage:** 123.0x

**Sequencing technology:** PacBio Sequel

IDs: 4536931 [UID] 13654338 [GenBank] 13759658 [RefSeq]

See [Genome](#) Information for **Strongylocentrotus purpuratus**

6. Scroll down to the bottom of the page to the section titled: “Global assembly definition”

Assembly Definition
Assembly Statistics

### Global assembly definition

[Download the full sequence report](#)

Click on the table row to see sequence details in the table to the right

The primary assembly unit does not have any assembled chromosomes or linkage groups. Please download the full sequence report for information on the scaffolds.

Assembly Unit Name
<b>Primary Assembly</b>
non-nuclear

- Click the “non-nuclear option.” For animals, look to see if there is a mitochondria option. For plants, look to see if there is a chloroplast option. Click the underlined link under “GenBank ID.” If there are not any choices, go back to step 5 and choose a different assembly. If there are no other assemblies available, repeat steps 3 and 4, following the lineage further back to check for assemblies in distantly related taxonomic groups.

Assembly Definition
Assembly Statistics

**Global assembly definition** [Download the full sequence report](#)

Click on the table row to see sequence details in the table to the right

Assembly Unit Name	Assembly Unit
Primary Assembly	
<b>non-nuclear</b>	

Assembly Unit: non-nuclear (GCF\_000028735.1)

Molecule name	GenBank ID	RefSeq ID
Mitochondrion MT	<a href="#">X12631.1</a>	= <a href="#">NC_001453.1</a>

- Shown below is the information for the assembled *Strongylocentrotus purpuratus* mitochondrial genome. If you scroll down, you will see all the genes in the assembly listed. For animals, try to find a COI gene. For plants, try to find a RuBisCO gene. You might have to skim through the genes or keyword search different terms. For example, the COI gene could be under “cytochrome c” or “COI.”

GenBank
Send to:
Change region shown
Customize view

**Strongylocentrotus purpuratus complete mitochondrial genome**

GenBank: X12631.1

[FASTA](#) [Graphics](#)

[Go to:](#)

LOCUS	X12631	15650 bp	DNA	circular INV 25-JUL-2016
DEFINITION	Strongylocentrotus purpuratus complete mitochondrial genome.			
ACCESSION	X12631			
VERSION	X12631.1			
KEYWORDS	12S ribosomal RNA; 16S ribosomal RNA; ATPase 6; ATPase 8; cytochrome b; cytochrome oxidase subunit 1; cytochrome oxidase subunit 2; cytochrome oxidase subunit 3; genome; mitochondrial DNA; NADH dehydrogenase subunit 1; NADH dehydrogenase subunit 2; NADH dehydrogenase subunit 3; NADH dehydrogenase subunit 4; NADH dehydrogenase subunit 4L; NADH dehydrogenase subunit 5; NADH dehydrogenase subunit 6; transfer RNA-Ala; transfer RNA-Arg; transfer RNA-Asn; transfer RNA-Asp; transfer RNA-Cys; transfer RNA-Gln; transfer RNA-Glu; transfer RNA-Gly; transfer RNA-His; transfer RNA-Ile; transfer RNA-Leu; transfer RNA-Lys; transfer RNA-Met; transfer RNA-Phe; transfer RNA-Pro; transfer RNA-Ser; transfer RNA-Thr; transfer RNA-Trp; transfer RNA-Tyr; transfer RNA-Val.			
SOURCE	mitochondrion Strongylocentrotus purpuratus (purple sea urchin)			
ORGANISM	<a href="#">Strongylocentrotus purpuratus</a>			

**Analyze this sequence**

Run BLAST

Pick Primers

Highlight Sequence Features

Find in this Sequence

**Related information**

Protein

PubMed

Taxonomy

Full text in PMC

Gene

Identical RefSeq

PubMed (Weighted)

9. Once you've found the gene you want, click on the underlined "gene" or "CDS" link. CDS stands for coding sequence. Clicking this link will show you where your sequence is in the genome assembly.

[CDS](#)  
.....

5785..7338  
/codon\_start=1  
/transl\_table=9  
/product="cytochrome oxidase subunit 1"  
/protein\_id="CAA31153.1"  
/db\_xref="GOA:P15544"  
/db\_xref="InterPro:IPR000883"  
/db\_xref="UniProtKB/Swiss-Prot:P15544"  
/translation="MQLSRWLFSTN HKDIGTLYLIFGAWAGMVGTAMSVIIRAELAQPGSLLNDDQIYNVVVTAHALVMIFFMVPMIMIGGFGNWLIPLMIGAPDMAFFRMNMSFWLIPPSFILLASAGVENAGTGTWIIYPPLSSNITHAGSSVDLAIFSLHLGASSILGLINFITTIINMRTPGMSLDRLPLFVWSVFVTAFLLLLSLPVLGAIITMLLTDRNINTTFFDPAGGGDPILFQHLFWLFGHPEVYIILIPGFGMISHVIAHYSKREPFYGLGLVYAMIAIGVLGFLVWAHHMFTVGMVDVTRAYFTAATMIIAVPTGLKVSWMAKLQGSNLQWSLPLLWTLGIVFLFTLGGLTGIVLANSSIDFVLHDTYYVVAHFHYVLSMGAVFAIFAGFTHWFPLFSGYSLHPLWGKVHFFIMFVGVNLTFFPQHFLGLAGMPRRYSYDPDAYTLWNTISSIGSTISVVAMLFFFLIWEAFASQREGITPEFSHASLEWQYTSFPPSHHTFDETPSTIIIVK"

10. On the bottom right, click the "Display: FASTA" option. This will take you to a new screen with the selected sequence in FASTA format.

5761 aacgtaaaga cttcaattaa aaacatgcaa ctaagacgat gattattttc tactaacacg  
5821 aaggacatcg gaacacttta tttattttt ggggcctgag ctggcatggt aggcacagct  
5881 atgagtgtga ttatccgtgc cgaattggca caacctgggt cctgtctaaa agatgacacg  
5941 atatacaag ttgtcgttac cgaactgagc ctatgctatg tttttctcat ggtaatgcca  
6001 aatatgattg gtggtatttg gaattgactc attocactaa tgatcggtgc gccagatagc  
6061 gcttcctccc gcatgaaaaa tatgattttt tgacttatto cctctcttt tatattactt  
6121 tttagcttcg caggagtaga aaaaaggaga ggaactggct gaactatcta cctctctctc  
6181 tctagtataa taacacacgc cggtaggtgc gtgatttagt caatctcttc ccttcacctg  
6241 gcgggtgctt cttccatctt gggctcattt aaattttata caacatttat taatatcgcc  
6301 acacgggaga tgtctttgga tctctctctt ttatctgtct gatcgtctt tgcactgccc  
6361 tcttctgccc tectttctct tccagtatta gcaaggacaa ttcaatgct tctcacagat  
6421 gytatataaa aaacacattt ctgtgacctt gcaaggaggg gatatccat tctatttttc  
6481 caactattct ggttttttgg acaccccgaa gttatatttc ttacttacc gpgatttggg  
6541 ttgatctaac acgttatagc tcaactactt gtaagcagag agctcttcgg atacctgggg  
6601 ttggtttatg ccatgattgc aataggagtt tttaggtacc ttgtctgggc ccacatatac  
6661 ttacagttag gcatggatgt tgatacacga gcatacttcc ctgcgcacac aatgattatt  
6721 gctgtcccaa caggattaaa ggttttcaaa tgaatggcaa agtcccaagg gtcatactta  
6781 caatgaagac tccctttatt atgaaccttg gggattgtat ttttattcac attaggagga  
6841 ctcaacagta ttgttttgc caattctccc atgactttg ttcttcatga taactactac  
6901 ttggtagctc actttcacta cgtttcttca atggggctg tatttgaat ctctcgtggt  
6961 ttaactcaat gttttccctt cttctctggt tatagctcac acccatatg aggaaggtt  
7021 caattcttca taatgtttgt tggagtcaac ttaacctttt tccctcaaca cttcttaggt  
7081 ttaggccgaa tgcacagagc gtaactagac tatccagagc cctatacatt ttgaataact  
7141 tctctctcaa ttgatcaac catctctgta gtggtatgac tatttttctt cttcttatac  
7201 tggatgacct tcatctctca acggtatgta atctcccaa atttctcaa cgtctcactc  
7261 taatgacaaat aaacctctgt tccpctcttc caaccaacct tcatgaaaac accctctaac  
7321 taattatttg taagttagt gtacgttaaa gagttatgtt aaggagaacc tctgattctg  
7381 gctcagatgg ttttggttca accccaatac ctttgaatct gccctgotta tagttatttt  
7441 atcaatgttt tacttaggac ttatgggaat ccttttaaaa cgactccatt tccatctcat  
7501 tctattatgc cttgagctac tctttatttc ctgttttata ggaattgaga ttgaaaaaaa  
7561 aaaaacaggg gtccctcaaa aaacaacatt caactttatc gttttgactc tagtcgcttg

5785..7338  
/codon\_start=1  
/transl\_table= 9  
/product="cytochrome oxidase subunit 1"  
/protein\_id=" CAA31153.1 "  
/db\_xref="GOA: P15544 "  
/db\_xref="InterPro: IPR000883 "  
/db\_xref="UniProtKB/Swiss-Prot: P15544 "  
/translation="MQLSRWLFSTN HKDIGTLYLIFGAWAGMVGTAMSVIIRAELAQPGSLLNDDQIYNVVVTAHALVMIFFMVPMIMIGGFGNWLIPLMIGAPDMAFFRMNMSFWLIPPSFILLASAGVENAGTGTWIIYPPLSSNITHAGSSVDLAIFSLHLGASSILGLINFITTIINMRTPGMSLDRLPLFVWSVFVTAFLLLLSLPVLGAIITMLLTDRNINTTFFDPAGGGDPILFQHLFWLFGHPEVYIILIPGFGMISHVIAHYSKREPFYGLGLVYAMIAIGVLGFLVWAHHMFTVGMVDVTRAYFTAATMIIAVPTGLKVSWMAKLQGSNLQWSLPLLWTLGIVFLFTLGGLTGIVLANSSIDFVLHDTYYVVAHFHYVLSMGAVFAIFAGFTHWFPLFSGYSLHPLWGKVHFFIMFVGVNLTFFPQHFLGLAGMPRRYSYDPDAYTLWNTISSIGSTISVVAMLFFFLIWEAFASQREGITPEFSHASLEWQYTSFPPSHHTFDETPSTIIIVK"

CDS Feature 3 of 12 X12631:1 segment Details Display: FASTA GenBank Help



FASTA ▾

## Strongylocentrotus purpuratus complete mitochondrial genome

GenBank: X12631.1

[GenBank](#) [Graphics](#)

>X12631.1:5785-7338 Strongylocentrotus purpuratus complete mitochondrial genome

```
ATGCAACTAAGACGATGATTATTTTCTACTAACCACAAGGACATCGGAACACTTTATTTAATTTTGGGG
CCTGAGCTGGCATGGTAGGCACAGCTATGAGTGTGATTATCCGTGCCGAGTTGGCACAACCTGGTTCGCT
GCTAAAAGATGACCAGATATACAAAGTGGTCGTTACCGCACATGCGCTAGTCATGATTTTCTTCATGGTA
ATGCCAATAATGATTGGTGGATTGGGAATTGACTCATTCCACTAATGATCGGTGCCCGAGATATGGCCT
TCCCCCGCATGAAAAATATGAGTTTTTGAAGTATTTCCCTTCTTTTATATTACTTTAGCGTCCGCAGG
AGTAGAAAAAGGAGCAGGAAGTGGCTGAAGTATCTACCTCCTCTCTCTAGTAAAATAACACACGCCGGT
AGGTCCGTTGATTTAGCAATCTTCTCCCTTCACCTGGCCGGTGCCTCTTCCATCTTGGGCCTCATTAAAT
TTATAACAACAATTATTAATATGCGGACACCGGGGATGTCTTTGGATCGTCTTCCCTTATTCGTCTGATC
CGTCTTTGTCACTGCCTTCTTGCTCCTCTTCTCTTCCAGTATTAGCAGGAGCAATTACAATGCTTCTC
ACAGATCGTAAAATAAACACAAGTCTTCTTGACCTGCAGGAGGGGGAGATCCAATTCTATTTCAACACC
TATTCTGGCTTTTTTGGACACCCCGAGGTGTATATTCTTATCTTACCGGGATTTGGTATGATCTCACACGT
TATAGCTCACTACTCTGGTAAGCGAGAGCCTTTTCGGATACCTGGGGTTGGTTTATGCCATGATTGCAATA
GGAGTTTTAGGATTCTTGTCTGGGCCCACCATATGTTTACAGTAGGGATGGATGTTGATACACGAGCAT
ACTTCACTGCCGCCACAATGATTATTGCTGTCCCAACAGGATTAAAGGTTTTCAGATGAATGGCAAAGCT
CCAAGGGTCTAATCTACAATGAAGACTCCCTTTATATGAACCTTGGGGATTGTATTTTATTACATTA
GGAGGACTCACAGGTATTGTTCTTGCCAATTCCCTCCATTGACTTTGTTCTTCATGATACCTACTACGTGG
TAGCTCACTTTCACTACGTTCTTTCAATGGGGGCTGTATTTGCAATCTTCGCTGGTTTCACTCACTGGTT
TCCCTCTTCTCTGGTTATAGCCTACACCCATTATGAGGAAAGGTTCACTTCTTCATAATGTTTGTGGA
GTCAACTTAACCTTTTTTCCCTCAACACTTCTTAGGTCTAGCCGGAATGCCACGACGGTACTCAGACTATC
CAGACGCCTATACACTTTGAAATACTATCTCCTCAATTGGATCAACCATCTCCGTAGTGGCTATGCTATT
TTTCTCTTCTTAATCTGGGAGGCCTTCGCTTCTCAACGGGAAGGAATCACCCAGAGTTCTCACACGCC
TCACTAGAGTGACAATACACCTCCTTTCCCCCTTCTCACCACACCTTCGATGAAACACCCCTTACCATAA
TTATTGTAAAGTAA
```

11. Copy and paste the sequence into a text editor (either Sublime or Atom, do not use TextEdit). **Make sure to include the header that begins with the ">" symbol.**



The screenshot shows a text editor window with a dark theme. The title bar reads ">X12631.1:5785-7338 Strongylocentrotus purpuratus". The left sidebar shows "OPEN FILES" with the current file selected. The main editor area displays a FASTA sequence starting with ">X12631.1:5785-7338 Strongylocentrotus purpuratus complete mitochondrial genome" followed by 24 lines of DNA sequence. The cursor is at the end of line 24, which is "TTATTGTAAAGTAA".

```
>X12631.1:5785-7338 Strongylocentrotus purpuratus complete mitochondrial genome
1 ATGCAACTAAGACGATGATTATTTTCTACTAACCACAAGGACATCGGAACACTTTATTTAATTTTGGGG
2 CCTGAGCTGGCATGGTAGGCACAGCTATGAGTGTGATTATCCGTGCCGAGTTGGCACAACTGGTTCGCT
3 GCTAAAAGATGACCAGATATACAAAGTGGTCGTTACCGCACATGCGCTAGTCATGATTTTCTCATGGTA
4 ATGCCAATAATGATTGGTGGATTGGGAATTGACTCATTCCACTAATGATCGGTGCGCCAGATATGGCCT
5 TCCCCCGCATGAAAAATATGAGTTTTTGACTTATCCCCCTTCTTTTATATTACTTTTAGCGTCCGCAGG
6 AGTAGAAAAAGGAGCAGGAAGTGGCTGAACATCTACCCTCCTCTCTAGTAAAAATAACACACGCCGGT
7 AGGTCGCTTGATTAGCAATCTTCTCCCTTACCTGGCCGGTGCCTCTTCCATCTTGGGCCTCATTAAAT
8 TTATAACAACAATTATTAATATGCGGACACCGGGGATGTCTTTGGATCGTCTTCTTTATTCGTCTGATC
9 CGTCTTTGTCACTGCCTTCTTGTCTCCTCTTCTCTTCCAGTATTAGCAGGAGCAATTACAATGCTTCTC
10 ACAGATCGTAAAAATAACACAACCTTTCTTTGACCTGCAGGAGGGGGAGATCCAATTTCTATTTCAACACC
11 TATTCGGCTTTTGGACACCCGAGGTGTATTTCTTATCTTACCGGGATTTGGTATGATCTCACACGT
12 TATAGCTCACTACTCTGGTAAGCGAGAGCCTTTCCGATACCTGGGTTGGTTTATGCCATGATTGCAATA
13 GGAGTTTATGATTCTTGTCTGGGCCACCATATGTTACAGTAGGGATGGATGTTGATACACGAGCAT
14 ACTTCACTGCCGCCACAATGATTATTGCTGTCCCAACAGGATTAAGGTTTTAGATGAATGGCAAAGCT
15 CCAAGGCTCTAATCTACAATGAAGACTCCCTTTTATGAACTTGGGGATTGATTTTTATTACATTA
16 GGAGGACTCACAGGTATTGTTCTTGGCAATCTCCATTGACTTTGTTCTTCATGATACCTACTACGTGG
17 TAGCTCACTTTCACTACGTTCTTCAATGGGGGCTGTATTTGCAATCTTCGCTGGTTTCACTACTGGTT
18 TCCCTCTTCTCTGGTTATAGCCTACACCCATTATGAGGAAAGGTTCACTTCTTCATAATGTTTGTGGA
19 GTCAACTTAACCTTTTCCCTCAACACTTCTTAGGTCTAGCCGGAATGCCACGACGGTACTCAGACTATC
20 CAGACGCCTATACACTTTGAAATACTATCTCCTCAATTGGATCAACCATCTCCGTAGTGGCTATGCTATT
21 TTTCTTCTTAAATCTGGGAGGCCTTCGTTCTCAACGGGAAGGAATCACCCAGAGTTCTCACACGCC
22 TCACTAGAGTGACAATACACCTCTTCCCTTCTCACCACACCTTCGATGAAACACCTCTACCATAA
23
24 TTATTGTAAAGTAA
```

12. Save this file as seed\_mito.fasta or seed\_chloro.fasta. Remember where you saved it!

