

Living members of the order **Carnivora** are characterized by a particular arrangement of the cheek tooth dentition called the **Carnassial Pair**, abbreviated as **P⁴/M₁**, for the **fourth upper premolar** and first lower molar. When the jaw closes, the **P⁴** passes outside the **M₁**, so as to create a shearing surface that assists in the processing of the typical carnivore diet, flesh. The **P⁴/M₁** form has been variously modified in various lineages, including reduced dentition in the insectivorous aardwolf (*Proteles cristata*) or crab eater seals (*Lobodon carcinophaga*), or as a massive crushing surface in hyaenas (*Crocuta*). Other taxa with less dependence on flesh have more or less flattened molariform dentition.

Living Carnivora occur in two distinct lineages, the dog-like Caniformia / Canoidea and the cat-like Feliformia / Feloidea. The former classically comprises seven families: **Canidae** (dogs), **Ursidae** (bears), **Mustelidae** (weasels & relatives), **Procyonidae** (raccoons & relatives), and the three marine forms, **Phocidae** (true or earless seals), **Otariidae** (eared seals), and **Odobenidae** (walrus). The latter classically comprises three forms, **Felidae** (cats), **Hyaenidae** (hyaenas), and **Viverridae** (mongooses & civets). The last family was the least understood, and has undergone extensive revision in the light of molecular evidence. The Felidae as well have been controversial, as to whether (seeming) morphological uniformity is reflected in single genera of small (*Felis*) and large (*Panthera*), or some other arrangement. [Kelt & Patton (2020) provides a compact overview, Wilson & Reeder (2005) an exhaustive review].

In this lab, we will apply three phylogenetic approaches to molecular data, in this case the complete set of 13 **mitochondrial DNA (mtDNA)** protein **coding regions** (ca. 12 kbp) from multiple species of carnivores. The methods are those available through the program **MEGA** (Molecular Evolutionary Genetics Analysis, v.12): **Neighbor Joining (NJ)**, **Maximum Parsimony (MP)**, and **Maximum Likelihood (ML)**.

The data are arranged in a **MEGA** file of **13 protein coding regions** from multiple carnivore species, aligned and spaced so as to preserve the first **5'-3' reading frame** of the first locus (**ND1**) on the **Heavy Strand** consistent through the end of the last (**Cytb**). This includes *reverse complements* of four loci that occur on the *Light Strand*.

We will use these data and methods to explore several classic and contemporary problems of Carnivore systematics and taxonomy. Each exercise will involve a subset of the sequence in the base file, plus several sequences to be retrieved from the **GenBank** library of taxonomic DNA sequence data. The data will be imported into the **MEGA** files as above, and the analyses done so as to investigate the questions

Methods

1. First learn how to retrieve sequences from **GenBank**. Go to the GenBank link to **Metazoa** in the lecture notes. Proceed down to the root of the **Carnivora** tree. Tick the box '*genomes*' and click on '*Go*'. The update screen will indicate the presence of a **complete mtDNA genome** and (or) a whole-genome project, as '*I*'.
 - a. Click on the '*I*': A window with 'See also 1 organelle- and plasmid-only records matching your search' should open. [OR, search for **NC_042596** or another sequence as instructed].
 - b. Click on that link. An Organism Overview will be presented, with a RefSeq link to a sequence file, typically beginning 'NC' and followed by six digits. Click the link.
 - c. This will take you to a Reference File for the mtDNA genome of interest, annotated in GenBank format with coded descriptions of the genome contents. In the white panel to the upper right, click on 'send to'. In the drop down menu, choose '**Coding Sequences**' as a FASTA file, then 'Create File'
 - d. A new 'Downloads' menu will open and ask what you want to do. Click on 'Save as', which will bring up a Windows box with 'sequence.txt'. SAVE AS a file name with a code like **NC123456** and the genus and species name of the genome source. SPELL CAREFULLY. Make sure it goes to the correct folder.
 - e. Proceed as above to add any additional GenBank records as required.
 - f. *Example:* At the root, click on the '*I*' for **Conepatus**, then click on 'see', then on '**NC_042596**'. Proceed as in (d) and download '**NC042596 Conepatus chinga.txt**'
2. Edit the file '**NC042596 Conepatus chinga.txt**' as indicated on the last three sheets. This requires removal of intervening text commentary.

3. To align the new sequence with those already aligned, open the **MEGA** datafile as directed. This is labelled something like **4250_Carnivore.mas** where *.mas is the alignment format.
 - a. Choose **Analyze**. Choose **Vertebrate mtDNA** as the genetic code, if not already chosen.
 - b. Move the slider all the way to the left. Notice that positions 1-3 are the **ATG** start **triplet** for the first coding locus, **ND1**. [This is *not* a *codon*, which occur *only* in **mRNA**]
 - c. Scan the slider forward to position **2014** for the **3'** end of **ND1**. There is a **Stop** triplet, followed by a spacer of exactly 4 x 3 b blank triplets. **ND2** begins with an **ATG** triplet at position **2026**.
 - i. Scanning forward through the rest of the file will show that all pre-existing sequences have been aligned so as to keep all loci constant in the first reading frame. Gaps are not consistently 12 or other multiples of three, for various reasons.
 - d. Import your new files, edited as in Step 2, as follows
 - i. Under the second **MEGA** pull-down menu Edit, choose 'Insert sequence from file'
 - ii. Choose the new sequence to be inserted & aligned.
 - iii. The new file will show up among the existing files. Move it next to a close relative species.
 - iv. Scan along up to position 2014: it should be apparent that homologous ND1 sequences among carnivores are typically quite similar, with major differences between caniform and feliform species.
 - v. In most cases, the end of the new sequence will terminate at position 2013: it may be necessary to pad the end with an 'a' or 'ag' to conform to the other sequences [mtDNA Stop triplets are frequently formed during polyadenylation of the **mRNA** with a string of 'aaaaa's, corresponding to a terminal 'g' on the Sense Strand]. Indicates edits as lower-case letters.
 - vi. In most cases, the beginnings of successive **Open Reading Frames** (ORFs) will begin in about the same spot, but may shift. Sequences are subject to **insertion – deletion** (*indel*) events with respect to other related sequences
 - vii. In all cases, micro-tune the alignments between your new sequences data and the old. This will typically involve insertion of a space (type a hyphen) or removal of one (place cursor on NEXT letter, backspace). There are numerous editing tricks that can be learned as your go on; the instructor will indicate some of these.
 1. Notably: the DNA data can be displaced as one-letter amino acid translations: this often makes adjustment of homologous region easier.
 - viii. **SAVE the intermediate results at frequent intervals:** MEGA is buggy and crashes a lot.

Systematics of Carnivora

Export the *.mas alignment to a *.meg file for phylogenetic analysis.

- e. Analysis can also be done within the *.mas file, but annotation and organization of files is more flexible and annotatable in *.meg. The drawback is that you have to go back to *.mas to modify the file content.
- f. Under 'Data' / 'Export alignment' / 'MEGA format'
 - i. If you're lucky, this will go smoothly.
 - ii. If not, a formatting error can hang you up. Consult instructors.
- g. You will have choice of **NJ**, **MP**, and **ML** options, with various choices of parameters within these modes. **Consult the course notes** or instructors for various options.
- h. Initial trees can be obtained with, corroboration from **Bootstrap** analysis (see lecture). These trees can be arranged, drawn, & exported as *.tiff figures for incorporation in your lab reports. Pretty pictures require practice.
- i. **Relationships of Giant Panda and Lesser Panda** to other caniform taxa. The Giant Panda (*Ailuropoda melanoleuca*) differs from typical bears in several respects, including a diet that comprises exclusively bamboo, highly modified cheek teeth and jaw structure consistent with this diet, and a biogeographic distribution limited to continental China. The Giant Panda also lacks specific skull characteristics otherwise characteristic of bears. This, along with its biogeographic and dietary similarities to the Lesser or Red Panda (*Proailurus fulgens*) have historically been taken to suggest that *Ailuropoda* and *Proailurus* are closer relatives of the Procyonidae than of Ursidae. Test this hypothesis by an appropriate selection of ursid, procyonid, and other caniform sequences, along with those of *Ailuropoda* and *Proailurus*.
- j. **Genetic differentiation of feliform and caniform morphology.** The occurrence of only three feliform families versus seven caniform families has been taken to indicate greater morphological divergence in the latter, and to predict (as in the comparison of Great Apes) greater molecular differentiation. Test this hypothesis by measuring differences among representative species for each family. Choose the **type** genus where possible, ie, *Canis* from Canidae, *Felis* from Felidae, and so on.
- k. **Molecular diversity within the Felidae.** In comparison to other carnivore families, it is often stated that all felids are pretty much alike, the only notable exception being the cheetah (*Acionyx jubata*), which has non-retractile claws. Otherwise, cats have often been divided between two genera, the smaller (*Felis*) and the larger (*Panthera*). Test this hypothesis by an appropriate selection of felid species, including *Acionyx* and a variety of larger and smaller species.
- l. **Molecular & taxonomic diversity within mongooses, civets, & relatives.** Unlike carnivore families with Nearctic and Palearctic distributions, relationships among the Global South and African species of catlike carnivores known as mongooses and civets (among other local names) were poorly understood until the advent of molecular data. Textbooks treated them as a single family, Viverridae. Test the hypothesis of a single monophyletic family. Suggest an alternative arrangement of viverrid taxa, and their relationship to the other two feliform families, Felidae and Hyaenidae.
- m. **Evolution of pagophilic behavior in Phocidae.** Based on a small region of the mtDNA **Cytochrome b gene**, [Perry et al. \(1995\)](#) proposed a theory of the origins of pagophilic ("ice-loving") breeding behavior in earless seals of the North Atlantic. Test this hypothesis with the complete mitogenomic sequences for all species in that study.

1) NC042596 *Conepatus chinga.txt* file looks like:

```
>lcl|NC_042596.1_cds_YP_009646897.1_1 [gene=ND1] [locus_tag=FI654_mgp13] [db_xref=GeneID:40412674]
[protein=NADH dehydrogenase subunit 1] [transl_except=(pos:955..955,aa:TERM)]
[protein_id=YP_009646897.1] [location=2758..3712] [gbkey=CDS]
ATGCTAATTATTAACATTATTTCTCTAATCGTACCTATTCTCTGGCTGTAGCTTTCTTGACCCCTAGTAG
AACGTAAAGTACTAGGTTATATACAGCTTCGAAAAGGTCCTAACATTGTAGGACCTTACGGACTACTTCA
ACCCATCGCAGACGCTGTAAACTATTACAAAAGAACCTACGACCCCTACATCTCTATATCCATA
TTTATTTTAGCCCCATCTCTGCCCTAACACTAGCCCTAACAAATATGAATTCCACTGCCAATACCATACC
CACTTACCAACATAAACCTGAGTATTCTCTCATACTAGCCATATCAAGCCTCGCTGTTTACTCAATTCT
ATGATCAGGTTGATCTTCAAACCTCAAATATGCTCTAATTGGAGCCTTACGAGCTGTAGCCCAACAATC
TCCTACGAAGTAACACTAGCCATTATCTTATTATCAGTATTATTAATAAACGGATCATTCACTCTATCTA
CTTTAATCACTACACAAGAACACTTATGATTAATCTTCCCCGCATGACCACTGGCCATAATATGATTTCAT
CTCAACCCCTAGCAGAACTAATCGAGCACCATTGATTAAACAGAAGGAGAATCAGAACTAGTATCAGGA
TTTAATGTAGAATACGAGCTGGCCCTTCGCACTATTCTTTTAGCAGAAATATGTAACATCATTATAA
TAAATATCCTTACTACCATTTCTATTCTCGGCGCATATCATAATCCATTAAATGCCGGAACATATTTCAAT
TAACCTACTACTAAAACACTACTCCTAACAAATTTCTTTCTGTGAATCCGAGCATCATACCCCGATTTC
CGCTATGATCAACTTATACACTTATTATGAAAAAATTTTTTACCCTGACACTAGCCCTTTGCATATGAC
ACATAGCCCTACCTATTATCAGCAAGCATCCCTCCACAAACAT
>lcl|NC_042596.1_cds_YP_009646898.1_2 [gene=ND2] [locus_tag=FI654_mgp12] [db_xref=GeneID:40412677]
[protein=NADH dehydrogenase subunit 2] [transl_except=(pos:1042..1042,aa:TERM)]
[protein_id=YP_009646898.1] [location=3922..4963] [gbkey=CDS]
ATAAAACCCACTATCATCAATCGTTATCTCATCAACCGTTATCATAGGAACCATAATAGTAATAAAGCT
CCCCTGATTATTAATCTGAATTGGATTGAAATAAATATACTAACAATTATCCCCATTTCTAATAAAAAA
TTTTAACCCACGAGCCATAGAAGCATCAACAAAATATTTTCTAACACAGGCTACAGCATCAATAATCCTA
ATACTAGGCATTGCCACAAATCTACTACTAACAGGACAATGAACAATGCTAAACATGCTAAACCCAATTG
CATCAAACATCACTACCATTTGCCCTAGCAATGAAACTCGGACTCTCCCTTTCCACTTCTGAGTACCCGA
AGTAACCCAAGGAGTCTCTATATTTTCAGGAACAGTACTACTAACCTGACAAAAAATCGCACCATTATCA
GTCTTATACCAAATTTTACCCTACCTAATTAACCTAGCCCTACTAACCATTCTAGCCATTATATCAGTACTAG
TAGGAGGATGAGGAGGACTAAACCAAACCTCAGCTCCGAAAAATCCTGGCTTACTCATCCATCTCCCATAT
AGGATGAATAACAGCAGTATTAACATATAATCCAACCTTAACAACCTTAACCTACTAATTTATATTATA
ATAACAACAGGAACCTTTTTATCTATTCTTAAATTTCTCCACCACCACACTATCACTATCTCACACAT
GAAACAAGTTACCCCTAAGCACACTCCTAGTACTGATACTCATACTATCCATCGGAGGCTTACCACCCCT
ATCAGGTTTTATACCAAAATGAATAATCATTCAGAAGCTCACAAAAACGACCTAATTATGATACCAACA
CTCATAGCAATAGCAGCCCTAGTAAATCTATATTTTTTACATACGACTTATATACACAACAACACTCACAA
TATTTCCCTCATCAAACAACATAAAAAATAAAATGACAGTTTCGAAAACCTCAAAAAATAATACCCCTTCTACC
ACCACTAATTATCTATCAACCATATCACTACCACTCTCCCAATACTAATAATTCTACACT
>lcl|NC_042596.1_cds_YP_009646899.1_3 [gene=COX1] [locus_tag=FI654_mgp11] [db_xref=GeneID:40412682]
[protein=cytochrome c oxidase subunit I] [protein_id=YP_009646899.1] [location=5354..6898]
[gbkey=CDS]
ATGTTTCATCAACCGATGGCTATTTTCTACAAATCACAAAGACATCGGCACTCTTTACCTCCTATTCGGGG
CATGAGCCGGAATAGCTGGAACAGCCCTTAGTCTTCTAATTCGAGCTGAATTAGGACAACCTGGAACCCCT
ATTAGGAGATGACCAGATTTACAATGTAGTAGTCACAGCCCATGCATTCATATAATCTTCTCATAGTA
ATACCAATCATAAATTGGAGGCTTTGGAAACTGACTAGTGCCATTAATGATTGGAGCCCCAGATATAGCAT
TTCCACGAATAAATAACATAAGCTTCTGATTATTACCCCATCCCTCCTTCTACTTATAGCTTCTCTCCAT
GGTAGAAGCAGGTGCGAGAACTGGGTGAACAGTATATCCTCCACTAGCAGGTAACCTAGCCACGCAGGA
GCATCCGTAGATCTAACAATCTTCTCCTTACATCTAGCAGGTGTATCCTCAATTTTAGGGGCTATTAACT
TCATTACTACAATTATTAATATAAAACCCCTGCAATATCCCAATATCAAACCTCCCTGTTCGTATGATC
TGTAATAATTACAGCAGTCTCTTACTCTTATCCCTACCAGTATTGGCAGCCGGCATTACTATATTACTC
ACAGATCGAAACCTAAATACAACCTTCTTTGACCCAGCTGGAGGAGGAGATCCAATCCTATATCAACACT
TGTTCTGATTCTTTGGACATCCAGAAGTCTATATTCTAATTTTACCAGGATTGGGTATAATTCACATAT
CGTAACCTTATTACTCAGGCACAAAAAGAACCATTTGGATATATGGGTATAGTATGAGCAATAATGTCTATT
GGCTTCTTAGGCTTTATTGTATGAGCCACCATATATTTACTGTAGGTATAGACGTAGATACAGAGCCT
ACTTTACCTCAGCTACCATAAATCATTTGCAATTTCCAACCTGGAGTAAAGTATTTAGTTGACTAGCTACTTT
GCACGGAGGCAATATCAAATGATCCCTGCCATCTATGAGCACTAGGCTTCATTTTTTTATTCACAGTT
GGTGGCCTTACAGGTATTGTATTATCCAATTTCTTCACTAGACATTGTACTTCATGATACGTACTATGTAG
TAGCCCACTTCCACTATGTGTTATCAATGGGAGCAGTATTTGCAATCATAGGAGGATTGTTTCATTGATT
CCCCTTATTTTCAGGTTATACACTCAATGATGCATGAGCAAAAAATCCACTTCACAATTATATTTGTAGGA
GTAAACATAACATTTTCCCACAACATTTTCTAGGTCTATCAGGAATACCCGACGTTATTTCAGATTACC
CAGACGCTTATACAACATGAAACACAGTATCTTCTATAGGCTCTTTTATCTCGCTAACAGCAGTAATACT
AATAATCTTTATAATCTGAGAAGCCTTCGCATCCAAACGAGAAGTTCTAGTAGTAAATTATACCACCACT
AATATTGAATGACTACATGGATGTCTCCCCCATATCACACATTTGAAGAACCCTTATGTAATACTAA
AATAA
```

And so on through thirteen blocks

2) Paste a twelve (12) hyphen string '-----' over every 1cl text block **except** the first;

```
>1cl|NC_042596.1_eds_YP_009646897.1_1 [gene=ND1] [locus_tag=FI654_mgp13] [db_xref=GeneID:40412674]
[protein=NADH dehydrogenase subunit 1] [transl_except=(pos:955..955,aa:TERM)]
[protein_id=YP_009646897.1] [location=2758..3712] [gbkey=CDS]-----
ATGCTAATTATTAAACATTATTTCTCTAATCGTACCTATTCTCCTGGCTGTAGCTTTCTTGACCCCTAGTAG
AACGTAAAGTACTAGGTTATATACAGCTTCGAAAAGGTCCTAACATTGTAGGACCTTACGGACTACTTCA
ACCCATCGCAGACGCTGTAAACTATTACAAAAGAACCCTACGACCCCTCACATCCTCTATATCCATA
TTTATTTTAGCCCCATCTCTGCCCTAACACTAGCCCTAACAAATATGAATTCCTACTGCCAATACCATACC
CACTTACCAACATAAACTGAGTATTCTCTTCACTAGCCATATCAAGCCTCGCTGTTTACTCAATTCT
ATGATCAGGTTGATCTTCAAACCTCAAATATGCTCTAATTGGAGCCTTACGAGCTGTAGCCCAAACAATC
TCCTACGAAGTAACACTAGCCATTATCTTATTCAGTATTATTAATAAACGGATCATTCACTCTATCTA
CTTTAATCACTACACAAGAACACTTATGATTAATCTTCCCCGCATGACCACTGGCCATAATATGATTCAT
CTCAACCCCTAGCAGAACTAATCGAGCACCATTGATTAAACAGAAGGAGAATCAGAACTAGTATCAGGA
TTTAATGTAGAATACGAGCTGGCCCTTCGCACTATTCTTTTAGCAGAATATGCTAACATCATTATAA
TAAATATCCTTACTACCATTCTATTCTTCGGCGCATATCATAATCCATTAAATGCCGGAACCTATATTCAAT
TAACCTTCACTACTTAAACACTACTCCTTAACAATTTCTTTCCTGTGAATCCGAGCATCATACCCCGATT
CGCTATGATCAACTTATACACTTATTATGAAAAAATTTTACCCTTGACACTAGCCCTTTCATATGAC
ACATAGCCCTACCTATTATCACAGCAAGCATCCCTCCACAACAT----->1cl|NC_042596.1_eds_YP_009646898.1_2 [gene=ND2]
[locus_tag=FI654_mgp12] [db_xref=GeneID:40412677]
[protein=NADH dehydrogenase subunit 2] [transl_except=(pos:1042..1042,aa:TERM)]
[protein_id=YP_009646898.1] [location=3922..4963] [gbkey=CDS]-----
ATAAAACCCACTATCATATCGTTATCTCATCAACCGTTATCATAGGAACCATAATAGTAATAATAAGCT
CCCACTGATTATTAATCTGAATTGGATTGAAATAAATATACTAACAATTATCCCCATTCTAATAAAAAA
TTTTAACCCACGAGCCATAGAAGCATCAACAAAATATTTTCTAACACAGGCTACAGCATCAATAATCCTA
ATACTAGGCATTGCCACAAATCTACTACTAACAGGACAATGAACAATGCTAAACATGCTAAACCCAATTG
CATCAACATCACTACCATTGCCCTAGCAATGAAACTCGGACTCTCCCTTTCCACTTCTGAGTACCCGA
AGTAACCCAAGGAGTCTCTATATTTTCAGAACAGTACTACTAACCTGACAAAAAATCGCACCATTATCA
GCTTTATACCAAATTTTACCATTCCATTAAACCTAGCCTACTAACCATTCTAGCCATTATATCAGTACTAG
TAGGAGGATGAGGAGGACTAAACCAACTCAGCTCCGAAAAATCCTGGCTTACTCATCCATCTCCCATAT
AGGATGAATAACAGCAGTATTAACATATAATCCAACCTTAACAACCTTAACCTACTAATTTATATTATA
ATAACAACAGGAACCTTTTTATTTACTTTATTCTAAATTTCTCCACCACCACACTATCACTATCTCACACAT
GAAACAAGTTACCCCTAAGCACACTCCTAGTACTGATACTCATACTATCCATCGGAGGCCTACCACCCCT
ATCAGGTTTTATACCAAATGAATAATCATTCAGAAGCTCAGAAAAACGACCTAATATGATACCAACA
CTCATAGCAATAGCAGCCCTACTAAATCTATATTTTACATACGACTTATATACACAACAACACTCACAA
TATTCCCTCATCAAACAACATAAAAAATAAATGACAGTTCGAAAACCTCAAAAATAATACCCCTTCTACC
ACCACTAATTATTTCTATCAACCATATCACTACCCTCTCCCAATACTAATAATTCACACT-----
>1cl|NC_042596.1_eds_YP_009646899.1_3 [gene=COX1] [locus_tag=FI654_mgp11] [db_xref=GeneID:40412682]
[protein=cytochrome c oxidase subunit I] [protein_id=YP_009646899.1] [location=5354..6898]
[gbkey=CDS]-----
ATGTTTCATCAACCGATGGCTATTTTCTACAAATCACAAAGACATCGGCACTCTTTACCTCCTATTCGGGG
CATGAGCCGGAATAGCTGGAACAGCCCTTAGTCTTCTAATTGAGCTGAATTAGGACAACCTGGAACCTT
ATTAGGAGATGACCAGATTTACAATGTAGTAGTCACAGCCCATGCATTCATATAATCTTCTTCATAGTA
ATACCAATCATAAATTGGAGGCTTTGGAAACTGACTAGTGCCATTAATGATTGGAGCCCCAGATATAGCAT
TTCCACGAATAAATAACATAAGCTTCTGATTATTAACCCCATCCCTCCTTCTACTTATAGCTTCCTCCAT
GGTAGAAGCAGGTGCAGGAACGGGTGAACAGTATATCCTCCACTAGCAGGTAACCTAGCCCCACGCAGGA
GCATCCGTAGATCTAACAATCTTCTCCTTACATCTAGCAGGTGTATCCTCAATTTTAGGGGCTATTAAC
TCATTACTACAATATTAAATATAAAACCCCTGCAATATCCCAATATCAAACCTCCCTGTTCGTATGATC
TGTAATAATTACAGCAGCTCTTCTTACTCTTATCCCTACCAGTATTGGCAGCCGGCATTACTATATTACTC
ACAGATCGAAACCTAAATACAACCTTCTTTGACCCAGCTGGAGGAGGAGATCCAATCCTATATCAACACT
TGTTCTGATTCTTTGGACATCCAGAAGTCTATATTCTAATTTTACCAGGATTGGGTATAATTTACATAT
CGTAACCTATTACTCAGGCAAAAAAGAACCTTTGGATATATGGGTATAGTATGAGCAATAATGTCTATT
GGCTTCTTAGGCTTTATTGTATGAGCCACCATATATTACTGTAGGTATAGACGTAGATACAGGACCT
ACTTTACCTCAGCTACCATAATCATTGCAATTTCAACTGGAGTAAAGTATTTAGTTGACTAGCTACTTT
GCACGGAGGCAATATCAAATGATCCCTGCCATACTATGAGCACTAGGCTTCATTTTTTTATTTCACAGTT
GGTGGCCTTACAGGTATTGTATTATCCAATTTCTTACTAGACATTGTACTTCATGATACGTACTATGTAG
TAGCCCACTTCCACTATGTGTTATCAATGGGAGCAGTATTGCAATCATAGGAGGATTGTTCATTGATT
CCCCTTATTTTCAGGTTATACACTCAATGATGCATGAGCAAAAATCCACTTCACAATTATATTGTAGGA
GTAAACATAACATTTTCCCACAACATTTTCTAGGTCTATCAGGAATACCCCGACGTATTTCAGATTACC
CAGACGCTTATACAACATGAACACAGTATCTTCTATAGGCTCTTTTATCTCGCTAACAGCAGTAATACT
AATAATCTTTATAATCTGAGAAGCCTTCGCATCCAAACGAGAAGTTCTAGTAGTAAATTATACCACCCT
AATATTGAATGACTACATGGATGTCTCCCCCATATCACACATTTGAAGAACCCACTTATGTAATACTAA
AATAA-----
```

And so on through thirteen blocks

3) Copy species name from original file. Text file now look like this. Save edited file as 'NC042596
Conepatus chinga.txt'

```
>lel+NC_042596 Conepatus chinga
ATGCTAATTATTAAACATTATTTCTCTAATCGTACCTATTCTCTGGCTGTAGCTTTCTTGACCCTAGTAG
AACGTAAAGTACTAGGTTATATACAGCTTCGAAAAGGTCCTAACATTGTAGGACCTTACGGACTACTTCA
ACCCATCGCAGACGCTGTAAACTATTACAAAAGAACCTACGACCCCTCACATCCTCTATATCCATA
TTTATTTTAGCCCCCATCTTGCCCTAACACTAGCCCTAACAAATATGAATTCCTACTGCCAATACCATACC
CACTTACCAACATAAACCTGAGTATTCTCTTCATACCTAGCCATATCAAGCCTCGCTGTTTACTCAATTCT
ATGATCAGGTTGATCTTCAAACCTCAAATATGCTCTAATTGGAGCCTTACGAGCTGTAGCCCAAACAATC
TCCTACGAAGTAACACTAGCCATTATCTTATTATCAGTATTATTAATAAACGATCATTCACTCTATCTA
CTTTAATCACTACACAAGAACACTTATGATTAATCTTCCCCGCATGACCACTGGCCATAATATGATTCAT
CTCAACCCCTAGCAGAACTAATCGAGCACCATTGATTTAACAGAAGGAGAATCAGAACTAGTATCAGGA
TTTAATGTAGAATACGAGCTGGCCCCCTTCGCACTATTCTTTTAGCAGAATATGCTAACATCATTATAA
TAAATATCCTTACTACCATTCTATTCTTCGGCGCATATCATAATCCATTAATGCCGGAACCTATATCAAT
TAACCTCACTACTAAAACACTACTCCTAACAAATTTCTTCTGTGAATCCGAGCATCATACCCCGATTTC
CGCTATGATCAACTTATACACTTATTATGAAAAAATTTTACCCTGACACTAGCCCTTTGCATATGAC
ACATAGCCCTACCTATTATCACAGCAAGCATCCCTCCACAAACAT
```

```
-----
ATAAAACCCACTATCATAATCGTTATCTCATCAACCGTTATCATAGGAACCATAATAGTAATAATAAGCT
CCCCTGATTATTAATCTGAATTGGATTGAAATAAATATACTAACAAATTATCCCATTCTAATAAAAAA
TTTTAACCCACGAGCCATAGAAGCATCAACAAATATTTCTAACACAGGCTACAGCATCAATAATCCTA
ATACTAGGCATTGCCACAAATCTACTACTAACAGGACAATGAACAATGCTAAACATGCTAAACCCAATTG
CATCAACATCACTACCATTGCCCTAGCAATGAACTCGGACTCTCCCCTTTCCACTTCTGAGTACCCGA
AGTAACCAAGGAGTCTCTATATTTTCAAGAACAGTACTACTAACCTGACAAAAAATCGCACCATTATCA
GTCTTATACCAAAATTTACCATTCCATTAACCTAGCCTACTAACCTTCTAGCCATTATATCAGTACTAG
TAGGAGGATGAGGAGGACTAAACCAACTCAGCTCCGAAAAATCCTGGCTTACTCATCCATCTCCCATAT
AGGATGAATAACAGCAGTATTAAACATATAATCCAACCTTAACAACCTTAACCTACTAATTTATATTATA
ATAACAACAGGAACCTTTTATTACTTATTCTAAATTTCTTCCACCACCACTATCACTATCTCACACAT
GAAACAAGTTACCCCTAAGCACACTCTAGTACTGATACTCATACTATCCATCGGAGGCTTACCACCCCT
ATCAGGTTTTATACCAAAATGAATAATCATTCAAGAACTCACAAAAACGACCTAATTATGATACCAACA
CTCATAGCAATAGCAGCCCTACTAAATCTATATTTTACATACGACTTATATACACAACAACACTCACAA
TATTTCCCTCATCAACAACATAAAAAATAAATGACAGTTTCGAAAACTCAAAAAATAATACCCCTTCTACC
ACCACTAATTATTCTATCAACCATATCACTACCCTCTCCCAATACTAATAATTCTACACT
```

```
-----
ATGTTTCATCAACCGATGGCTATTTTCTACAAATCACAAAGACATCGGCACTCTTTACCTCCTATTCGGGG
CATGAGCCGGAATAGCTGGAACAGCCCTTAGCTTCTAATTGAGCTGAATTAGGACAACCTGGAACCCCT
ATTAGGAGATGACCAGATTTACAATGTAGTAGTCACAGCCCATGCATTTCATATAATCTTCTTCATAGTA
ATACCAATCATAATTGGAGGCTTTGGAACTGACTAGTGCCATTAATGATTGGAGCCCCAGATATAGCAT
TTCCACGAATAAATAACATAAGCTTCTGATTATTACCCCATCCCTCCTTCTACTTATAGCTTCCCTCAT
GGTAGAAGCAGGTGCAGGAACGGGTGAACAGTATATCCTCCACTAGCAGGTAACCTAGCCCACGCAGGA
GCATCCGTAGATCTAACAACTTCTCTCTTACATCTAGCAGGTGTATCTCAATTTTAGGGGCTATTAAT
TCATTACTACAATATTAAATATAAAACCCCTGCAATATCCCAATATCAAACCTCCCTGTTCGTATGATC
TGTAATAATTACAGCAGTCTCTTACTCTTATCCCTACCAGTATTGGCAGCCGGCATTACTATATTACTC
ACAGATCGAAACCTAAATACAACCTTCTTTGACCCAGCTGGAGGAGGAGATCCAATCCTATATCAACACT
TGTTCTGATTCTTTGGACATCCAGAAGTCTATATTCTAATTTTACCAGGATTTGGTATAATTTACATAT
CGTAACCTATTACTCAGGCAAAAAAGAACCTTTGGATATATGGGTATAGTATGAGCAATAATGTCTATT
GGCTTCTTAGGCTTTATTTGTATGAGCCACCATATATTACTGTAGGTATAGACGTAGATACACGAGCCT
ACTTTACCTCAGCTACCAATAATCATTGCAATTCCAACCTGGAGTAAAGTATTTAGTTGACTAGCTACTTT
GCACGGAGGCAATATCAAATGATCCCTGCCATACTATGAGCACTAGGCTTCATTTTTTTATTACAGTT
GGTGGCCTTACAGGTATTGTATTATCCAATTTCTCACTAGACATTGTACTTCATGATACGTACTATGTAG
TAGCCCACTTCCACTATGTGTTATCAATGGGAGCAGTATTTGCAATCATAGGAGGATTTGTTCAATTGATT
CCCCTTATTTTCAAGTTATACACTCAATGATGCATGAGCAAAAATCCACTTCACAATTATATTGTAGGA
GTAAACATAACATTTTCCCAACATTTTCTAGGTCTATCAGGAATACCCGACGTATTTCAGATTACC
CAGACGCTTATACAACATGAAACACAGTATCTTCTATAGGCTCTTTTATCTCGCTAACAGCAGTAATACT
AATAATCTTTATAATCTGAGAAGCCTTCGCATCCAAACGAGAAGTTCTAGTAGTAATTTATACCACCACT
AATATTGAATGACTACATGGATGTCCTCCCCCATATCACACATTTGAAGAACCCACTTATGTAATACTAA
AATAA
```

And so on for all thirteen blocks, as a continuous file.
Save as a text file with extension ".fas", and the accession
number and taxon name