wolf

November 21, 2025

1 Wolf diet analysis

1.1 Preparing the reference database

We'll use a small version of *Genbank* containing only mammal sequences shorter than 20kb.

```
obigrep -L 20000 -Z \
        ../course/data/Genbank/Release_264 \
        > gb264_small_mam.fasta.gz
```

On that small DB I'll run `obipcr`.

Vertebrate primers:

- forward: TTAGATACCCCACTATGC
- reverse: TAGAACAGGCTCCTCTAG

We'll allow for 4 mismatches at most on each primer.

[2]:
```
obipcr --forward TTAGATACCCCACTATGC \
       --reverse TAGAACAGGCTCCTCTAG \
       -L 200 -e 4 -Z \
       --no-progressbar \
       ../course/data/Genbank/Release_264/small.fasta.gz \
       > vert01_raw_db.fasta.gz
```

```
INFO[0000] Number of workers set 32
INFO[0000] Found 1 files to process
INFO[0000] ../course/data/Genbank/Release_264/small.fasta.gz mime type:
text/fasta
INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout
```

[3]:
```
obiuniq -m taxid -Z \
        vert01_raw_db.fasta.gz \
        > vert01_uniq_db.fasta.gz
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] vert01_raw_db.fasta.gz mime type: text/fasta
```

```
INFO[0000] Running dereplication on disk with 100 chunks
INFO[0000] Keep sigletons in the output
INFO[0000] Starting data splitting
                                                - Reading sequences (20678/-,
72999 it/s) [0s] - Splitting data set (22774/-, 78503 it/s) [0s]

INFO[0000] Data splitted over 100 batches
INFO[0000] End of the data splitting
INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_0.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 0/100 : 108 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_1.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 1/100 : 105 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_10.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 2/100 : 83 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_11.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 3/100 : 100 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_12.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 4/100 : 81 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_13.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 5/100 : 322 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_14.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 6/100 : 1022 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_15.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 7/100 : 156 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_16.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 8/100 : 304 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_17.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 9/100 : 368 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_18.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 10/100 : 148 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_19.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 11/100 : 526 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_2.fastx mime type:
```

```
text/fasta
INFO[0000] Start processing of batch 12/100 : 185 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_20.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 13/100 : 110 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_21.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 14/100 : 115 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_22.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 15/100 : 305 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_23.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 16/100 : 61 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_24.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 17/100 : 215 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_25.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 18/100 : 228 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_26.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 19/100 : 112 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_27.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 20/100 : 642 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_28.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 21/100 : 167 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_29.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 22/100 : 573 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_3.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 23/100 : 60 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_30.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 24/100 : 592 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_31.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 25/100 : 215 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_32.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 26/100 : 183 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_33.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 27/100 : 93 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_34.fastx mime type:
```

```
text/fasta
INFO[0000] Start processing of batch 28/100 : 129 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_35.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 29/100 : 173 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_36.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 30/100 : 160 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_37.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 31/100 : 127 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_38.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 32/100 : 108 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_39.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 33/100 : 200 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_4.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 34/100 : 112 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_40.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 35/100 : 167 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_41.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 36/100 : 249 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_42.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 37/100 : 138 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_43.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 38/100 : 235 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_44.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 39/100 : 135 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_45.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 40/100 : 451 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_46.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 41/100 : 87 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_47.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 42/100 : 177 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_48.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 43/100 : 166 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_49.fastx mime type:
```

```
text/fasta
INFO[0000] Start processing of batch 44/100 : 75 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_5.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 45/100 : 329 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_50.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 46/100 : 811 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_51.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 47/100 : 298 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_52.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 48/100 : 128 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_53.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 49/100 : 70 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_54.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 50/100 : 47 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_55.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 51/100 : 241 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_56.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 52/100 : 167 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_57.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 53/100 : 136 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_58.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 54/100 : 261 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_59.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 55/100 : 265 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_6.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 56/100 : 146 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_60.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 57/100 : 244 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_61.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 58/100 : 153 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_62.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 59/100 : 220 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_63.fastx mime type:
```

```
text/fasta
INFO[0000] Start processing of batch 60/100 : 299 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_64.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 61/100 : 850 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_65.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 62/100 : 147 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_66.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 63/100 : 132 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_67.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 64/100 : 216 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_68.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 65/100 : 105 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_69.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 66/100 : 857 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_7.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 67/100 : 87 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_70.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 68/100 : 180 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_71.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 69/100 : 181 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_72.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 70/100 : 94 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_73.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 71/100 : 98 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_74.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 72/100 : 49 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_75.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 73/100 : 264 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_76.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 74/100 : 367 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_77.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 75/100 : 133 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_78.fastx mime type:
```

```
text/fasta
INFO[0000] Start processing of batch 76/100 : 88 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_79.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 77/100 : 1445 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_8.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 78/100 : 1718 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_80.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 79/100 : 128 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_81.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 80/100 : 146 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_82.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 81/100 : 108 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_83.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 82/100 : 100 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_84.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 83/100 : 254 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_85.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 84/100 : 119 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_86.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 85/100 : 370 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_87.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 86/100 : 279 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_88.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 87/100 : 202 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_89.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 88/100 : 224 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_9.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 89/100 : 405 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_90.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 90/100 : 190 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_91.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 91/100 : 104 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_92.fastx mime type:
```

```
text/fasta
INFO[0000] Start processing of batch 92/100 : 96 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_93.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 93/100 : 58 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_94.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 94/100 : 79 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_95.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 95/100 : 222 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_96.fastx mime type:
text/fasta
INFO[0000] Start processing of batch 96/100 : 4158 sequences
INFO[0000] /tmp/obiseq_chunks_488015107/chunk_97.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 97/100 : 129 sequences
INFO[0001] /tmp/obiseq_chunks_488015107/chunk_98.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 98/100 : 183 sequences
INFO[0001] /tmp/obiseq_chunks_488015107/chunk_99.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 99/100 : 348 sequences
```

[4]:

```
obiannotate --add-lca-in taxid \
            -t ../course/data/ncbitaxo_20251118.tgz \
            vert01_uniq_db.fasta.gz \
  | obiannotate -t ../course/data/ncbitaxo_20251118.tgz \
            --taxonomic-rank -Z \
            > vert01_lca_db.fasta.gz
```

```
INFO[0001] NCBI Taxdump Tar Archive detected:
../course/data/ncbitaxo_20251118.tgz
INFO[0001] Loading Taxonomy nodes
INFO[0001] NCBI Taxdump Tar Archive detected:
../course/data/ncbitaxo_20251118.tgz
INFO[0001] Loading Taxonomy nodes
INFO[0016] 2706727 Taxonomy nodes read
INFO[0016] Loading Taxon names
INFO[0016] 2706727 Taxonomy nodes read
INFO[0016] Loading Taxon names
INFO[0038] 2706727 taxon names read
INFO[0038] Loading Merged taxa
INFO[0038] 2706727 taxon names read
INFO[0038] Loading Merged taxa
INFO[0038] 93509 merged taxa read
INFO[0038] Set as default taxonomy NCBI Taxonomy
INFO[0038] Number of workers set 16
```

```
INFO[0038] Found 1 files to process
INFO[0038] vert01_uniq_db.fasta.gz mime type: text/fasta
INFO[0038] On output use JSON headers
INFO[0038] Output is done on stdout
INFO[0038] Data is writen to stdout
WARN[0038] GU981038_sub[458..556]: Taxid 150090 has to be updated to
taxon:3371150 [Pseudosoriculus fumidus]@species
WARN[0038] KY410960_sub[517..615]: Taxid 1620840 has to be updated to
taxon:2740394 [Sundasciurus altitudinis]@species
WARN[0038] KX381710_sub[181..280]: Taxid 263451 has to be updated to
taxon:3370180 [Lophostoma silvicola]@species
WARN[0038] JF694095_sub[447..545]: Taxid 164531 has to be updated to
taxon:3370248 [Microperoryctes ornatus]@species
WARN[0038] AY495457_sub[449..548]: Taxid 258864 has to be updated to
taxon:3370479 [Ozimops planiceps]@species
WARN[0038] KY464180_sub[516..615]: Taxid 2021987 has to be updated to
taxon:3370147 [Lenothrix cana]@species
WARN[0038] KY581660_sub[517..616]: Taxid 478698 has to be updated to
taxon:3370289 [Mops plicatus]@species
WARN[0038] AY495540_sub[446..543]: Taxid 258960 has to be updated to
taxon:2720888 [Vespadelus darlingtoni]@species
WARN[0038] AJ430552_sub[441..540]: Taxid 410298 has to be updated to
taxon:3370000 [Gerbilliscus afer]@species
WARN[0038] MF038679_sub[446..543]: Taxid 2093327 has to be updated to
taxon:3369970 [Eptesicus dutertreus]@species
WARN[0038] JX520511_sub[433..531]: Taxid 35665 has to be updated to
taxon:3370058 [Heterogeomys hispidus]@species
WARN[0038] MN807618_sub[513..611]: Taxid 119565 has to be updated to
taxon:3371118 [Montemys delectorum]@species
WARN[0038] MK211397_sub[445..543]: Taxid 114227 has to be updated to
taxon:3370452 [Onychogalea frenata]@species
WARN[0038] MF038680_sub[440..537]: Taxid 258906 has to be updated to
taxon:3371012 [Cnephaeus hottentotus]@species
WARN[0038] MZ708834_sub[527..626]: Taxid 224955 has to be updated to
taxon:3370766 [Scaptonyx fusicauda]@species
WARN[0038] AJ389548_sub[455..553]: Taxid 35664 has to be updated to
taxon:3370057 [Heterogeomys heterodus]@species
WARN[0038] KX591624_sub[446..543]: Taxid 1898424 has to be updated to
taxon:2778565 [Laephotis stanleyi]@species
WARN[0038] AF203773_sub[449..548]: Taxid 110940 has to be updated to
taxon:3371137 [Paremballonura atrata]@species
WARN[0038] OR257582_sub[644..744]: Taxid 3055736 has to be updated to
taxon:3370154 [Leopardus pardinoides]@species
WARN[0039] JQ039218_sub[446..543]: Taxid 568927 has to be updated to
taxon:3369693 [Afronycteris nanus]@species
WARN[0038] OM334923_sub[449..547]: Taxid 209424 has to be updated to
taxon:3370523 [Petaurista grandis]@species
WARN[0039] JF694114_sub[450..548]: Taxid 1031278 has to be updated to
```

```
taxon:3370506 [Perameles notina]@species
WARN[0039] KX381233_sub[178..277]: Taxid 40227 has to be updated to
taxon:3369931 [Dermanura gnomus]@species
WARN[0039] KX381261_sub[181..280]: Taxid 263451 has to be updated to
taxon:3370180 [Lophostoma silvicola]@species
WARN[0039] KX754507_sub[450..549]: Taxid 224955 has to be updated to
taxon:3370766 [Scaptonyx fusicauda]@species
WARN[0039] JN255346_sub[448..547]: Taxid 190653 has to be updated to
taxon:3369760 [Cercartetus nana]@species
WARN[0039] AY530083_sub[459..558]: Taxid 264759 has to be updated to
taxon:3370991 [Atopogale cubana]@species
WARN[0039] OP328298_sub[837..934]: Taxid 169057 has to be updated to
taxon:3371009 [Cnephaeus bottae]@species
WARN[0039] HM561626_sub[498..595]: Taxid 59452 has to be updated to
taxon:3371021 [Cnephaeus serotinus]@species
WARN[0039] KP202255_sub[1281..1381]: Taxid 61412 has to be updated to
taxon:3369756 [Caracal auratus]@species
WARN[0039] AF263234_sub[447..543]: Taxid 124746 has to be updated to
taxon:3371123 [Neoeptesicus furinalis]@species
WARN[0039] KR537889_sub[522..620]: Taxid 61085 has to be updated to
taxon:3369802 [Crocidura beata]@species
WARN[0039] AY495466_sub[452..549]: Taxid 258906 has to be updated to
taxon:3371012 [Cnephaeus hottentotus]@species
WARN[0039] AY495514_sub[446..543]: Taxid 177184 has to be updated to
taxon:2778567 [Pseudoromicia brunnea]@species
WARN[0039] KX754505_sub[452..551]: Taxid 224955 has to be updated to
taxon:3370766 [Scaptonyx fusicauda]@species
WARN[0039] MH801935_sub[521..620]: Taxid 2588392 has to be updated to
taxon:3370481 [Parablarinella griselda]@species
WARN[0039] AF442092_sub[447..546]: Taxid 263451 has to be updated to
taxon:3370180 [Lophostoma silvicola]@species
WARN[0039] HM106321_sub[518..617]: Taxid 55048 has to be updated to
taxon:3370370 [Neogale frenata]@species
WARN[0039] AY495467_sub[447..544]: Taxid 59452 has to be updated to
taxon:3371021 [Cnephaeus serotinus]@species
WARN[0039] KX754504_sub[450..549]: Taxid 224955 has to be updated to
taxon:3370766 [Scaptonyx fusicauda]@species

INFO[0039] 93509 merged taxa read
INFO[0039] Set as default taxonomy NCBI Taxonomy
INFO[0039] Number of workers set 16
INFO[0039] Reading sequences from stdin in guessed
INFO[0039] - mime type: text/fasta
INFO[0039] On output use JSON headers
INFO[0039] Output is done on stdout
INFO[0039] Data is writen to stdout
```

[5]:

```
obicsv -k taxonomic_rank vert01_lca_db.fasta.gz \
  | tail -n +2 \
  | sort \
  | uniq -c
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] vert01_lca_db.fasta.gz mime type: text/fasta
```

```
      4 clade
     52 family
    374 genus
      6 no rank
      1 order
   3180 species
      1 species group
     46 subfamily
      6 subgenus
      1 suborder
    180 subspecies
      6 tribe
```

[6]:

```
obigrep --require-rank family \
        -t ../course/data/ncbitaxo_20251118.tgz \
        --update-taxid \
        vert01_lca_db.fasta.gz \
        > vert01_clean_db.fasta.gz
```

```
INFO[0000] NCBI Taxdump Tar Archive detected:
../course/data/ncbitaxo_20251118.tgz
INFO[0000] Loading Taxonomy nodes
INFO[0007] 2706727 Taxonomy nodes read
INFO[0007] Loading Taxon names
INFO[0016] 2706727 taxon names read
INFO[0016] Loading Merged taxa
INFO[0016] 93509 merged taxa read
INFO[0016] Set as default taxonomy NCBI Taxonomy
INFO[0016] Number of workers set 16
INFO[0016] Found 1 files to process
INFO[0016] vert01_lca_db.fasta.gz mime type: text/fasta
INFO[0016] On output use JSON headers
INFO[0016] Output is done on stdout
INFO[0016] Data is writen to stdout
```

[7]:

```
obicsv -k taxonomic_rank vert01_clean_db.fasta.gz \
  | tail -n +2 \
  | sort \
  | uniq -c
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] vert01_clean_db.fasta.gz mime type: text/fasta


     52 family
    374 genus
      6 no rank
   3180 species
      1 species group
     46 subfamily
      6 subgenus
    180 subspecies
      6 tribe
```

[8]:

```
obicsv -k taxid vert01_clean_db.fasta.gz \
  | tail -n +2 \
  | sort \
  | uniq \
  | awk -F'@' '{print $2}' \
  | sort \
  | uniq -c
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] vert01_clean_db.fasta.gz mime type: text/fasta


     21 family
    188 genus
      3 no rank
   2021 species
      1 species group
     23 subfamily
      3 subgenus
     98 subspecies
      2 tribe
```

[9]:

```
obiannotate --taxonomic-path \
            -t ../course/data/ncbitaxo_20251118.tgz \
            -Z \
            vert01_clean_db.fasta.gz \
```

```
            > vert01_ref_db.fasta.gz
```

```
INFO[0000] NCBI Taxdump Tar Archive detected:
../course/data/ncbitaxo_20251118.tgz
INFO[0000] Loading Taxonomy nodes
INFO[0007] 2706727 Taxonomy nodes read
INFO[0007] Loading Taxon names
INFO[0017] 2706727 taxon names read
INFO[0017] Loading Merged taxa
INFO[0017] 93509 merged taxa read
INFO[0017] Set as default taxonomy NCBI Taxonomy
INFO[0017] Number of workers set 16
INFO[0017] Found 1 files to process
INFO[0017] vert01_clean_db.fasta.gz mime type: text/fasta
INFO[0017] On output use JSON headers
INFO[0017] Output is done on stdout
INFO[0017] Data is writen to stdout
```

1.2 Analyzing the metabarcoding data

1.2.1 Step 1: Pairing the Reads

We'll use the `obipairing` *OBITools4* command.

[10]:
```
obipairing -F ../course/data/Wolf_diet/wolf_F.fastq \
           -R ../course/data/Wolf_diet/wolf_R.fastq \
           -Z \
           > wolf_paired.fastq.gz
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] ../course/data/Wolf_diet/wolf_F.fastq mime type: text/fastq
INFO[0000] Found 1 files to process
INFO[0000] ../course/data/Wolf_diet/wolf_R.fastq mime type: text/fastq
INFO[0000] Start of the sequence Pairing using 16 workers
INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout
INFO[0000] Initializing the DNA Scoring matrix
                                                - Reading sequences (15525/-,
59229 it/s) [0s] | Reading sequences (19407/-, 43953 it/s) [0s] | Reading
sequences (27170/-, 60156 it/s) [0s] / Reading sequences (23288/-, 42042 it/s)
[0s] - Reading sequences (31052/-, 47481 it/s) [0s] | Reading sequences
(38815/-, 47481 it/s) [0s] \ Reading sequences (31052/-, 27527 it/s) [1s] -
Reading sequences (38815/-, 27527 it/s) [1s] - Reading sequences (41393/-, 30190
it/s) [1s] \ Reading sequences (41393/-, 27527 it/s) [1s]
                                                | Reading sequences (45276/-,
```

```
30190 it/s) [1s]
INFO[0003] End of the sequence Pairing
```

[11]:

```
obigrep -a mode=alignment wolf_paired.fastq.gz \
 | obicsv -k ali_length -k score_norm \
 > wolf_paired_scores.csv
```

```
INFO[0000] Number of workers set 16
INFO[0000] Number of workers set 16
INFO[0000] Reading sequences from stdin in guessed
INFO[0000] Found 1 files to process
INFO[0000] wolf_paired.fastq.gz mime type: text/fastq
INFO[0000] mode alignment
INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout
/ Reading sequences (5073/-, 50093 it/s) [0s] INFO[0000] - mime type:
text/fastq
                                                Reading sequences (13538/-,
46268 it/s) [0s] | Reading sequences (15231/-, 37250 it/s) [0s] / Reading
sequences (18616/-, 31567 it/s) [0s]                                        |
Reading sequences (3344/-, 8180 it/s) [0s] \ Reading sequences (25387/-, 31567
it/s) [0s]                                        - Writing CSV (6692/-, 11124
it/s) [0s] | Reading sequences (27086/-, 31567 it/s) [0s] / Reading sequences
(30462/-, 31567 it/s) [0s] | Reading sequences (13397/-, 11125 it/s) [0s] |
Writing CSV (13397/-, 11124 it/s) [0s] - Reading sequences (33849/-, 30908 it/s)
[1s]                                                \ Reading sequences
(15071/-, 13860 it/s) [1s] \ Writing CSV (15071/-, 13859 it/s) [1s] | Reading
sequences (20101/-, 13860 it/s) [1s] | Writing CSV (20101/-, 13859 it/s) [1s] -
Reading sequences (41896/-, 30908 it/s) [1s]
                                               / Reading sequences (23458/-,
13860 it/s) [1s] / Writing CSV (23458/-, 13859 it/s) [1s] \ Reading sequences
(26804/-, 13860 it/s) [1s] \ Writing CSV (26804/-, 13859 it/s) [1s] | Reading
sequences (34097/-, 13860 it/s) [1s] | Writing CSV (34097/-, 13859 it/s) [1s]
```

[12]:

```
obigrep  -a mode=alignment wolf_paired.fastq.gz \
  | obigrep -p 'annotations.score_norm >= 0.96 &&
                annotations.ali_length > 55 &&
                annotations.ali_length < 65' \
           -Z \
           > wolf_paired_good.fastq.gz
```

```
INFO[0000] Number of workers set 16
INFO[0000] Number of workers set 16
INFO[0000] Reading sequences from stdin in guessed
INFO[0000] Found 1 files to process
INFO[0000] wolf_paired.fastq.gz mime type: text/fastq
```

```
INFO[0000] mode alignment
INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout
/ Reading sequences (3378/-, 17818 it/s) [0s] INFO[0000] - mime type:
text/fastq
INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout
                                                Reading sequences (8457/-, 28787
it/s) [0s] | Reading sequences (13538/-, 27278 it/s) [0s] - Reading sequences
(15228/-, 22019 it/s) [0s] / Reading sequences (3348/-, 6690 it/s) [0s] |
Reading sequences (18614/-, 22019 it/s) [0s] - Reading sequences (5021/-, 6690
it/s) [0s] - Reading sequences (23692/-, 22019 it/s) [1s] | Reading sequences
(27080/-, 21165 it/s) [1s] / Reading sequences (28768/-, 21165 it/s) [1s] \
Reading sequences (10044/-, 8925 it/s) [1s] - Reading sequences (35544/-, 21165
it/s) [1s] | Reading sequences (16749/-, 8925 it/s) [1s] | Reading sequences
(37232/-, 21165 it/s) [1s] - Reading sequences (18427/-, 8925 it/s) [1s] /
Reading sequences (42318/-, 24145 it/s) [1s] \ Reading sequences (43585/-, 24145
it/s) [1s] | Reading sequences (20101/-, 11693 it/s) [1s]
                                               / Reading sequences (23458/-,
11693 it/s) [1s] - Reading sequences (26803/-, 11693 it/s) [1s] \ Reading
sequences (31830/-, 11693 it/s) [1s]
```

[13]:

```
obicount wolf_paired.fastq.gz | csvlook
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_paired.fastq.gz mime type: text/fastq
                                                Reading sequences (16921/-,
47077 it/s) [0s] | Reading sequences (23692/-, 51190 it/s) [0s] / Reading
sequences (32155/-, 56006 it/s) [0s] - Reading sequences (40629/-, 56006 it/s)
[0s] \ Reading sequences (42321/-, 56006 it/s) [0s]
| entities |         n |
| -------- | --------- |
| variants |    45,276 |
| reads    |    45,276 |
| symbols  | 7,386,937 |
```

[14]:

```
obicount wolf_paired_good.fastq.gz | csvlook
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_paired_good.fastq.gz mime type: text/fastq
                                                Reading sequences (13765/-,
54484 it/s) [0s] \ Reading sequences (26237/-, 74257 it/s) [0s]
| entities |         n |
| -------- | --------- |
```

```
| variants |    27,955 |
| reads    |    27,955 |
| symbols  | 4,302,597 |
```

1.3 Extracting the barcode

Using the `obimultiplex`command.

[15]:

```
obimultiplex -s ../course/data/Wolf_diet/wolf_data_wolf_diet_ngsfilter.csv \
             -u wolf_unassign.fastq.gz \
             -Z \
             wolf_paired_good.fastq.gz \
             > wolf_assign.fastq.gz
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_paired_good.fastq.gz mime type: text/fastq
INFO[0000] Reading NGSFilter file:
../course/data/Wolf_diet/wolf_data_wolf_diet_ngsfilter.csv
INFO[0000] No BOM detected
INFO[0000] NGSFilter configuration mimetype: text/ngsfilter-csv
INFO[0000] 3 parameters found
INFO[0000] Read 5 records
INFO[0000] First record: [experiment sample sample_tag forward_primer
reverse_primer]
INFO[0000] Set tag matching mode to strict
INFO[0000] Set global allowed primer mismatches to 2
INFO[0000] Disallows indels for primer matching
INFO[0000] Unassigned sequences saved in file: wolf_unassign.fastq.gz
INFO[0000] Sequence demultiplexing using 16 workers
INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout
INFO[0000] On output use JSON headers
                                                Reading sequences (10324/-,
42547 it/s) [0s] | Reading sequences (12045/-, 27315 it/s) [0s] \ Reading
sequences (17208/-, 22941 it/s) [0s] - Reading sequences (22802/-, 22941 it/s)
[1s] \ Reading sequences (27955/-, 22941 it/s) [1s]
```

[16]:

```
obicount wolf_unassign.fastq.gz | csvlook
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_unassign.fastq.gz mime type: text/fastq

| entities |      n |
| -------- | ------ |
| variants |    243 |
```

```
| reads    |    243 |
| symbols  | 23,598 |
```

[17]:

```
obicount wolf_assign.fastq.gz | csvlook
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_assign.fastq.gz mime type: text/fastq
                                               eading sequences (4415/-, 14719
it/s) [0s] / Reading sequences (8829/-, 17607 it/s) [0s] - Reading sequences
(11478/-, 17607 it/s) [0s] \ Reading sequences (14125/-, 17607 it/s) [0s] /
Reading sequences (16774/-, 17607 it/s) [0s] - Reading sequences (19425/-, 19283
it/s) [1s] \ Reading sequences (23836/-, 19283 it/s) [1s]
| entities |         n |
| -------- | --------- |
| variants |    27,712 |
| reads    |    27,712 |
| symbols  | 2,580,858 |
```

[21]:

```
obicsv -k sample wolf_assign.fastq.gz \
  | tail -n +2 \
  | sort \
  | uniq -c
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_assign.fastq.gz mime type: text/fastq
                                                Reading sequences (4416/-, 18923
it/s) [0s] - Writing CSV (4416/-, 18918 it/s) [0s] | Reading sequences (8831/-,
20349 it/s) [0s] | Writing CSV (8831/-, 20347 it/s) [0s] / Reading sequences
(11477/-, 21376 it/s) [0s] / Writing CSV (11477/-, 21374 it/s) [0s] \ Reading
sequences (14125/-, 21376 it/s) [0s] \ Writing CSV (14125/-, 21374 it/s) [0s] |
Reading sequences (15008/-, 21376 it/s) [0s] | Writing CSV (15008/-, 21374 it/s)
[0s] - Reading sequences (19425/-, 21376 it/s) [1s] - Writing CSV (19425/-,
21374 it/s) [1s] \ Reading sequences (20308/-, 18061 it/s) [1s] \ Writing CSV
(20308/-, 18058 it/s) [1s] / Reading sequences (23836/-, 18061 it/s) [1s] /
Writing CSV (23836/-, 18058 it/s) [1s]

   6447 13a_F730603
   6066 15a_F730814
   9567 26a_F040644
   5632 29a_F260619
      1 sample
```

[19]:

```
obiuniq -m sample -Z \
        wolf_assign.fastq.gz \
        > wolf_uniq.fasta.gz
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_assign.fastq.gz mime type: text/fastq
INFO[0000] Running dereplication on disk with 100 chunks
INFO[0000] Keep sigletons in the output
INFO[0000] Starting data splitting
                                                 Reading sequences (4415/-,
16703 it/s) [0s] \ Splitting data set (4415/-, 12929 it/s) [0s] | Reading
sequences (6179/-, 14028 it/s) [0s] | Splitting data set (7062/-, 15883 it/s)
[0s] / Reading sequences (8829/-, 16246 it/s) [0s] - Splitting data set (8829/-,
13801 it/s) [0s] - Reading sequences (10595/-, 16246 it/s) [0s] \ Splitting data
set (12361/-, 13801 it/s) [0s] \ Reading sequences (15008/-, 16246 it/s) [0s] |
Splitting data set (15008/-, 13801 it/s) [0s] | Reading sequences (15893/-,
16246 it/s) [0s] - Splitting data set (18542/-, 13801 it/s) [1s] - Reading
sequences (19425/-, 18704 it/s) [1s] \ Splitting data set (20308/-, 17161 it/s)
[1s] \ Reading sequences (21191/-, 18704 it/s) [1s] / Splitting data set
(25604/-, 17161 it/s) [1s] / Reading sequences (26484/-, 18704 it/s) [1s]

INFO[0001] Data splitted over 100 batches
INFO[0001] End of the data splitting
INFO[0001] On output use JSON headers
INFO[0001] Output is done on stdout
INFO[0001] Data is writen to stdout
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_0.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 0/100 : 37 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_1.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 1/100 : 14 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_10.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 2/100 : 13 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_11.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 3/100 : 13 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_12.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 4/100 : 355 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_13.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 5/100 : 21 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_14.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 6/100 : 18 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_15.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 7/100 : 38 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_16.fastx mime type:
```

```
text/fasta
INFO[0001] Start processing of batch 8/100 : 17 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_17.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 9/100 : 11 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_18.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 10/100 : 7 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_19.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 11/100 : 20 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_2.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 12/100 : 25 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_20.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 13/100 : 7756 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_21.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 14/100 : 43 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_22.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 15/100 : 27 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_23.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 16/100 : 10 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_24.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 17/100 : 23 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_25.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 18/100 : 15 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_26.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 19/100 : 16 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_27.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 20/100 : 46 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_28.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 21/100 : 31 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_29.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 22/100 : 15 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_3.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 23/100 : 48 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_30.fastx mime type:
```

```
text/fasta
INFO[0001] Start processing of batch 24/100 : 46 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_31.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 25/100 : 63 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_32.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 26/100 : 31 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_33.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 27/100 : 16 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_34.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 28/100 : 38 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_35.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 29/100 : 42 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_36.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 30/100 : 163 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_37.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 31/100 : 15 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_38.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 32/100 : 40 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_39.fastx mime type:
text/fasta
INFO[0001] Start processing of batch 33/100 : 24 sequences
INFO[0001] /tmp/obiseq_chunks_2438688034/chunk_4.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 34/100 : 5825 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_40.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 35/100 : 20 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_41.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 36/100 : 23 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_42.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 37/100 : 97 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_43.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 38/100 : 27 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_44.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 39/100 : 1863 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_45.fastx mime type:
```

```
text/fasta
INFO[0002] Start processing of batch 40/100 : 42 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_46.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 41/100 : 40 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_47.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 42/100 : 65 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_48.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 43/100 : 39 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_49.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 44/100 : 43 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_5.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 45/100 : 38 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_50.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 46/100 : 72 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_51.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 47/100 : 12 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_52.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 48/100 : 15 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_53.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 49/100 : 75 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_54.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 50/100 : 42 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_55.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 51/100 : 39 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_56.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 52/100 : 30 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_57.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 53/100 : 28 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_58.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 54/100 : 40 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_59.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 55/100 : 32 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_6.fastx mime type:
```

```
text/fasta
INFO[0002] Start processing of batch 56/100 : 35 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_60.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 57/100 : 35 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_61.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 58/100 : 8405 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_62.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 59/100 : 29 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_63.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 60/100 : 25 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_64.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 61/100 : 79 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_65.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 62/100 : 27 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_66.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 63/100 : 25 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_67.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 64/100 : 21 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_68.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 65/100 : 23 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_69.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 66/100 : 21 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_7.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 67/100 : 48 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_70.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 68/100 : 15 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_71.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 69/100 : 26 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_72.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 70/100 : 29 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_73.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 71/100 : 13 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_74.fastx mime type:
```

```
text/fasta
INFO[0002] Start processing of batch 72/100 : 26 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_75.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 73/100 : 15 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_76.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 74/100 : 15 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_77.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 75/100 : 21 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_78.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 76/100 : 57 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_79.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 77/100 : 18 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_8.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 78/100 : 21 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_80.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 79/100 : 17 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_81.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 80/100 : 15 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_82.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 81/100 : 55 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_83.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 82/100 : 30 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_84.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 83/100 : 47 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_85.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 84/100 : 37 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_86.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 85/100 : 10 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_87.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 86/100 : 21 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_88.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 87/100 : 26 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_89.fastx mime type:
```

```
text/fasta
INFO[0002] Start processing of batch 88/100 : 9 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_9.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 89/100 : 47 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_90.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 90/100 : 42 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_91.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 91/100 : 13 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_92.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 92/100 : 36 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_93.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 93/100 : 26 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_94.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 94/100 : 235 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_95.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 95/100 : 29 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_96.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 96/100 : 304 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_97.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 97/100 : 29 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_98.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 98/100 : 18 sequences
INFO[0002] /tmp/obiseq_chunks_2438688034/chunk_99.fastx mime type:
text/fasta
INFO[0002] Start processing of batch 99/100 : 33 sequences
```

[20]:
```
obicount wolf_uniq.fasta.gz | csvlook
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_uniq.fasta.gz mime type: text/fasta

| entities |      n |
| -------- | ------ |
| variants |    989 |
| reads    | 27,712 |
| symbols  | 97,427 |
```

1.4 Dataset cleaning

First step: looking at singleton sequences

[22]:

```
obicsv -k count wolf_uniq.fasta.gz \
  | sort -n \
  | uniq -c \
  | head -n 10
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_uniq.fasta.gz mime type: text/fasta


      1 count
    543 1
    125 2
     86 3
     66 4
     36 5
     14 6
     36 7
     19 8
     10 9
```

[23]:

```
obigrep -c 2 -Z \
        wolf_uniq.fasta.gz \
        > wolf_nosingleton.fasta.gz
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_uniq.fasta.gz mime type: text/fasta
INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout
```

[25]:

```
obicount wolf_nosingleton.fasta.gz | csvlook
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_nosingleton.fasta.gz mime type: text/fasta

| entities |      n |
| -------- | ------ |
| variants |    446 |
| reads    | 27,169 |
| symbols  | 43,760 |
```

Second step: Look at the sequence length distribution.

[26]:

```
obiannotate --length \
            wolf_nosingleton.fasta.gz \
| obicsv -k seq_length \
| sort -n \
| uniq -c
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_nosingleton.fasta.gz mime type: text/fasta
INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout
INFO[0000] Number of workers set 16
INFO[0000] Reading sequences from stdin in guessed

INFO[0000] - mime type: text/fasta


      1 seq_length
      5 4
      1 5
      1 8
    179 99
    259 100
      1 106
```

[27]:

```
obigrep -l 50 -Z \
        wolf_nosingleton.fasta.gz \
        > wolf_noshort.fasta.gz
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_nosingleton.fasta.gz mime type: text/fasta
INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout
```

[29]:

```
obicount wolf_noshort.fasta.gz | csvlook
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_noshort.fasta.gz mime type: text/fasta

| entities |      n |
| -------- | ------ |
```

```
| variants |    439 |
| reads    | 25,290 |
| symbols  | 43,727 |
```

Step 3: Look at ambiguous nucleotides.

```
[30]: obigrep -v -s '^[acgt]+$' \
              wolf_noshort.fasta.gz \
      | obicount | csvlook
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] Number of workers set 16
INFO[0000] Reading sequences from stdin in guessed
INFO[0000] wolf_noshort.fasta.gz mime type: text/fasta
INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout

INFO[0000] - mime type: text/fasta

| entities |   n |
| -------- | --- |
| variants |   4 |
| reads    |  10 |
| symbols  | 399 |
```

```
[31]: obigrep -Z -s '^[acgt]+$' \
              wolf_noshort.fasta.gz \
              > wolf_acgt.fasta.gz
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_noshort.fasta.gz mime type: text/fasta
INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout
```

```
[32]: obicount wolf_acgt.fasta.gz | csvlook
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_acgt.fasta.gz mime type: text/fasta

| entities |      n |
| -------- | ------ |
| variants |    435 |
| reads    | 25,280 |
```

```
| symbols  | 43,328 |
```

1.4.1 Running obiclean

Evaluating the ration threshold

[35]:
```
obiclean --save-ratio wolf_ratio.csv \
         --save-graph wolf_graph \
         wolf_acgt.fasta.gz \
         > wolf_obiclean_1.fasta
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_acgt.fasta.gz mime type: text/fasta

INFO[0000] Sequence dataset of 435 seqeuences loaded
INFO[0000] Dataset composed of 4 samples
                                                                       ror
graph]  44% |           | (18389952 it/s) [0s:0s][One error graph]  59%
|           | (17964160 it/s) [0s:0s][One error graph] 100%
|        | (19531142 it/s)[Annotate sequence status]  50% |
| (786 it/s) [0s:0s][Annotate sequence status]  75% |          | (1141
it/s) [0s:0s][Annotate sequence status] 100% |        | (1467 it/s)[Save
GML Graph files]  50% |            | (1060 it/s) [0s:0s][Save GML Graph
files]  75% |          | (1202 it/s) [0s:0s][Save GML Graph files] 100%
|        | (1212 it/s)[Save CSV stat ratio file]   8% |               |
(81360 it/s) [0s:0s][Save CSV stat ratio file]  12% |              | (74772
it/s) [0s:0s][Save CSV stat ratio file]  16% |              | (73066 it/s)
[0s:0s][Save CSV stat ratio file]  20% |              | (72208 it/s)
[0s:0s][Save CSV stat ratio file]  24% |              | (70560 it/s)
[0s:0s][Save CSV stat ratio file]  28% |             | (71283 it/s)
[0s:0s][Save CSV stat ratio file]  32% |             | (66951 it/s)
[0s:0s][Save CSV stat ratio file]  36% |             | (30132 it/s)
[0s:0s][Save CSV stat ratio file]  40% |            | (30654 it/s)
[0s:0s][Save CSV stat ratio file]  44% |            | (31945 it/s)
[0s:0s][Save CSV stat ratio file]  48% |           | (33099 it/s)
[0s:0s][Save CSV stat ratio file]  52% |           | (34555 it/s)
[0s:0s][Save CSV stat ratio file]  56% |           | (35574 it/s)
[0s:0s][Save CSV stat ratio file]  60% |          | (34914 it/s)
[0s:0s][Save CSV stat ratio file]  64% |          | (36083 it/s)
[0s:0s][Save CSV stat ratio file]  68% |          | (36119 it/s)
[0s:0s][Save CSV stat ratio file]  72% |          | (36945 it/s)
[0s:0s][Save CSV stat ratio file]  76% |         | (37940 it/s)
[0s:0s][Save CSV stat ratio file]  80% |         | (38694 it/s)
[0s:0s][Save CSV stat ratio file]  84% |         | (39511 it/s)
[0s:0s][Save CSV stat ratio file]  88% |        | (39952 it/s)
[0s:0s][Save CSV stat ratio file]  92% |        | (38738 it/s)
[0s:0s][Save CSV stat ratio file]  96% |       | (38713 it/s)
[0s:0s][Save CSV stat ratio file] 100% |       | (39438
```

```
it/s)INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout
```

[38]:

```
obiclean --detect-chimera \
         -r 0.1 -H \
         wolf_acgt.fasta.gz \
         > wolf_obiclean_2.fasta
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_acgt.fasta.gz mime type: text/fasta

INFO[0000] Sequence dataset of 435 seqeuences loaded
INFO[0000] Dataset composed of 4 samples
ph]  44% |            | (11275210 it/s) [0s:0s][One error graph]  59%
|           | (10893174 it/s) [0s:0s][One error graph] 100%
|        | (12325209 it/s)[Filter graph on abundance ratio]  50% |
| (54237 it/s) [0s:0s][Filter graph on abundance ratio]  75% |          |
(47337 it/s) [0s:0s][Filter graph on abundance ratio] 100% |         |
(44017 it/s)[Annotate sequence status]  50% |            | (18476 it/s)
[0s:0s][Annotate sequence status]  75% |          | (15161 it/s)
[0s:0s][Annotate sequence status] 100% |        | (3483
it/s)INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout
```

[39]:

```
obicount wolf_obiclean_2.fasta | csvlook
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_obiclean_2.fasta mime type: text/fasta

| entities |      n |
| -------- | ------ |
| variants |     30 |
| reads    | 22,608 |
| symbols  |  2,987 |
```

1.5 Taxonomical assignment

using `obitag`

[41]:

```
obitag -R ./vert01_ref_db.fasta.gz \
       --save-db ./vert01_ref_db_indexed.fasta \
       wolf_obiclean_2.fasta \
       > wolf_taxon.fasta
```

```
INFO[0000] Number of workers set 32
INFO[0000] Found 1 files to process
INFO[0000] wolf_obiclean_2.fasta mime type: text/fasta
INFO[0000] ./vert01_ref_db.fasta.gz mime type: text/fasta
INFO[0000] Set as default taxonomy taxon
/ Reading sequences (1/-, 5 it/s) [0s]
INFO[0000] 3851 reference sequences conserved on 3851
INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout
```

[45]:

```
obiannotate --number wolf_taxon.fasta \
| obiannotat e --set-identifier 'sprintf("MOTU_%03d", annotations.seq_number)' \
      > wolf_short_id.fasta
```

```
INFO[0000] Number of workers set 16
INFO[0000] Number of workers set 16
INFO[0000] Reading sequences from stdin in guessed
INFO[0000] Found 1 files to process
INFO[0000] wolf_taxon.fasta mime type: text/fasta
INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout

INFO[0000] - mime type: text/fasta
INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout
```

[48]:

```
obigrep -p 'max(annotations.obiclean_weight) >= 100' \
        wolf_short_id.fasta \
        > wolf_no_rare.fasta
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_short_id.fasta mime type: text/fasta
INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout
```

[49]:

```
obiannotate -k count \
            -k merged_sample \
            -k obiclean_weight \
            -k obitag_bestmatch \
            -k obitag_bestid \
            -k obitag_rank \
```

```
            -k taxid \
            wolf_no_rare.fasta \
            > wolf_taxon_cleaned.fasta
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_no_rare.fasta mime type: text/fasta
INFO[0000] On output use JSON headers
INFO[0000] Output is done on stdout
INFO[0000] Data is writen to stdout
```

[50]:

```
obicsv -i -s --auto \
       wolf_taxon_cleaned.fasta \
       > wolf_taxon_cleaned.csv
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_taxon_cleaned.fasta mime type: text/fasta
```

[51]:

```
obimatrix --transpose \
          -k id \
          -k taxid \
          -k obitag_bestid \
          --map obiclean_weight \
          wolf_no_rare.fasta | csvlook
```

```
INFO[0000] Number of workers set 16
INFO[0000] Found 1 files to process
INFO[0000] wolf_no_rare.fasta mime type: text/fasta
```

```
| taxon                                    | id       | obitag_bestid |
13a_F730603 | 15a_F730814 | 26a_F040644 | 29a_F260619 |
| ---------------------------------------- | -------- | ------------- |
----------- | ----------- | ----------- | ----------- |
| taxon:9611 [Canis]@genus                 | MOTU_066 |        1.000… |
9 |           4 |         328 |           1 |
| taxon:9992 [Marmota]@genus               | MOTU_006 |        0.990… |
0 |           0 |       8,744 |           0 |
| taxon:35500 [Pecora]@infraorder          | MOTU_014 |        0.950… |
0 |           0 |           0 |         152 |
| taxon:9860 [Cervus elaphus]@species      | MOTU_017 |        1.000… |
6,192 |           0 |           0 |           0 |
| taxon:55153 [Sciuridae]@family           | MOTU_020 |        0.949… |
0 |           0 |         146 |           0 |
| taxon:9858 [Capreolus capreolus]@species | MOTU_039 |        1.000… |
```

```
0 |       5,975 |           0 |       3,404 |
|                                           |          |               |
|             |             |             |
```

```
[ ]:
```