# **Supplementary Material**

#### ***SERPINA12* and *SLURP1* coding status assessment**

We applied Pseudo*Checker* (Alves et al., 2020) to assess the coding status of *SERPINA12* and *SLURP1* genes in 29 cetaceans (10 Mysticeti and 19 Odontoceti) and *H. amphibius*, the latter representing the closest relative to extant cetaceans, working as a pivotal species for estimating the time of gene loss events (Supplementary Tables 1 and 2, see below). Pseudo*Checker* is an integrated online pipeline capable of detecting cross-species conserved and non-conserved ORF-disruptive mutations (frameshift mutations, in-frame premature stop codons), greater levels of gene erosion (exon loss and full gene loss), only requiring, per analysis (per gene being tested), a single reference CDS plus the respective exon nucleotide sequence(s) and, for each target species, a corresponding orthologous genomic sequence against which the reference coding exons will be aligned, predicting the in-study gene CDS, further screened for inactivating evidence (Alves et al., 2020).

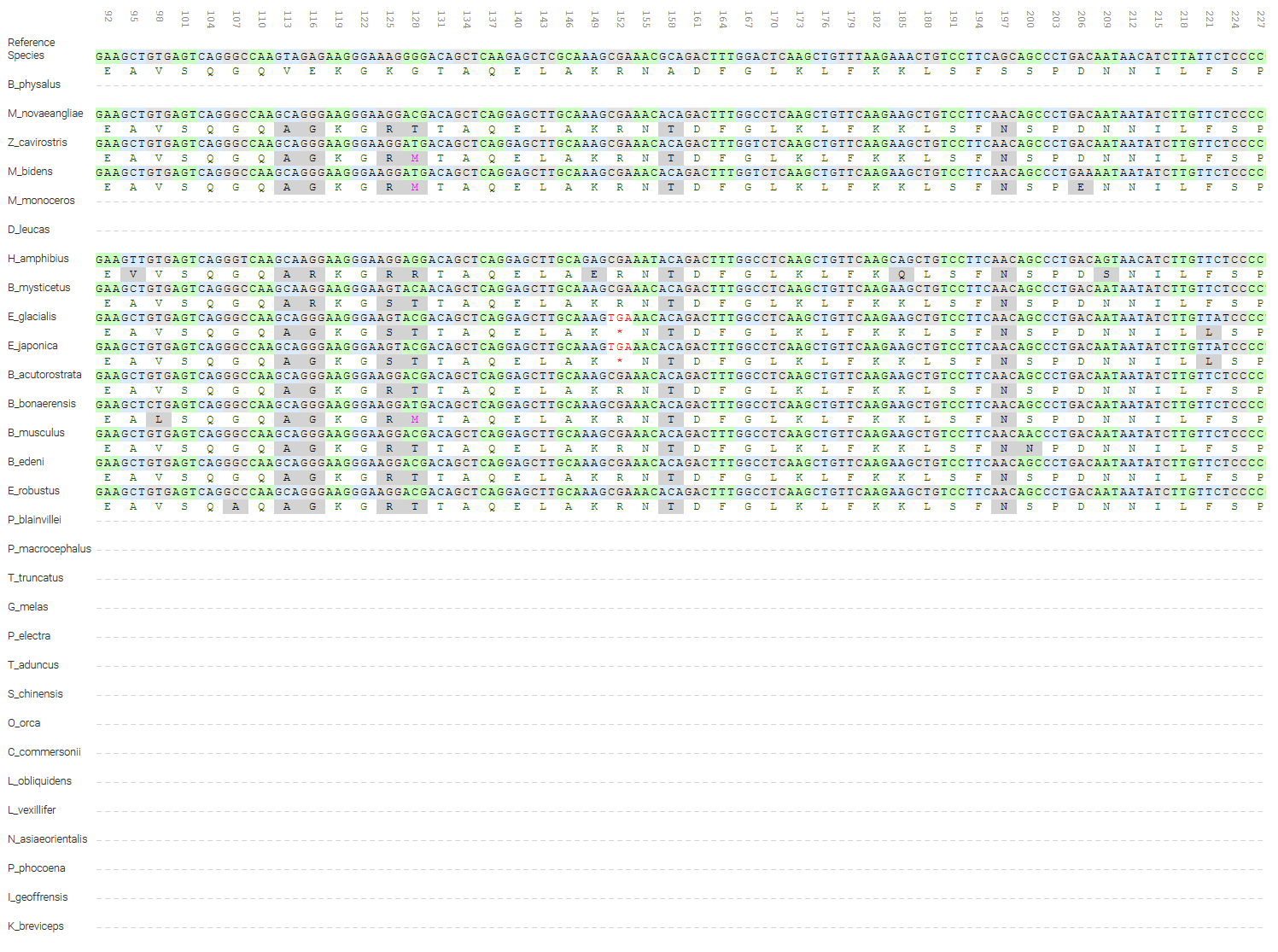
To this effect, for the two executed analyses, one regarding *SERPINA12* and the other concerning *SLURP1*, we used the *B. taurus* gene orthologs as a reference, and as target species corresponding genomic sequences, we extracted previously (annotated and unannotated) sequences from different biological databases (Supplementary Tables 1 and 2). In detail, (*i*) for species with annotated genome and annotated gene of interest at NCBI (i.e. *Balaenoptera acutorostrata scammoni*), we directly imported the corresponding genomic sequence into Pseudo*Checker*; (*ii*) for species with annotated genome but without annotation for the gene of interest at NCBI (i.e*. Tursiops truncatus*, *Delphinapterus leucas*), we submitted the respective genomic sequence flanked by the two gene of interest’s neighbouring genes in cetaceans, previously inferred while analysing the cases in (i) (*SERPINA4* and *SERPINA11* for *SERPINA12*; *LY6D* and *THEM6* for SLURP1); (*iii*) finally, for *Balaena mysticetus* (www.bowhead-whale.org), for species without annotated genome at NCBI (i.e. *Megaptera novaengliae*, *Phocoena phocoena*), and for species with available genome at DNA Zoo (www.dnazoo.org) (i.e. *Balaenoptera edeni*), blast searches were conducted against their genomes using, as a query, a set of three sequences including the *Bos taurus* CDS of the gene of interest, as well as the CDS’s of the neighbouring genes of it in the same species (*SERPINA4* and *SERPINA11* for *SERPINA12*; *LYPD2* and *THEM6* for SLURP1), further downloading the genomic sequence corresponding to the consensus blast hit across the hits obtained per each query sequence. Finally, through each Pseudo*Checker*’s analysis results, per species, identified disruptive mutations (when existing) were recovered. The visual demonstration of the detected ORF-disrupting mutations is represented through the Supplementary Figures 1 – 5 (see below Supplementary Tables 1 and 2).

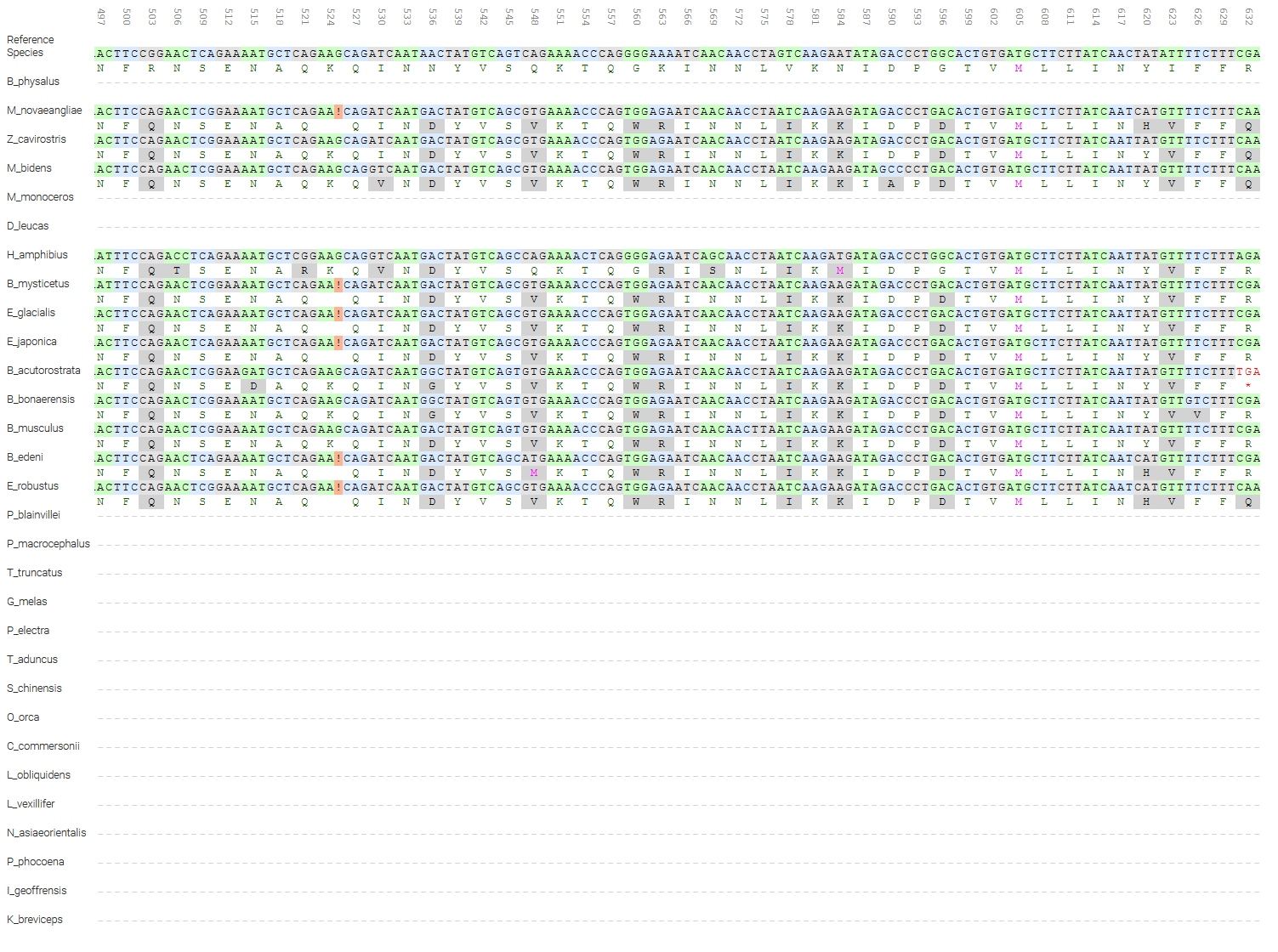
**Supplementary Table 1:** Analyzed genomic sequences regarding the Pseudo*Checker*’s coding status assessment of *SERPINA12* in cetaceans and *Hippopotamus amphibius*.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| # | Species | Family | Database | Genomic Sequence ID |
| 1 | *Tursiops aduncus* | Delphinidae | NCBI | NCQN01000009.1 |
| 2 | *Tursiops truncatus* | NCBI | NC\_047035.1 |
| 3 | *Sousa chinensis* | NCBI | RWJT01021752.1 |
| 4 | *Globicephala melas* | NCBI | NW\_022134922.1 |
| 5 | *Peponocephala electra* | DNA ZOO | HiC\_scaffold\_6 |
| 6 | *Cephalorhynchus commersonii* | DNA ZOO | HiC\_scaffold\_5 |
| 7 | *Lagenorhynchus obliquidens* | NCBI | NW\_020837965.1 |
| 8 | *Orcinus orca* | NCBI | NW\_004438415.1 |
| 9 | *Neophocaena asiaeorientalis asiaeorientalis* | Phocoenidae | NCBI | NW\_020172178.1 |
| 10 | *Phocoena phocoena* | NCBI | PKGA01135346.1 |
| 11 | *Monodon monoceros* | Monodontidae | NCBI | NW\_021703778.1 |
| 12 | *Delphinapterus leucas* | NCBI | NW\_022098016.1 |
| 13 | *Pontoporia blainvillei* | Pontoporiidae | NCBI | PGGH01071410.1 |
| 14 | *Inia geoffrensis* | Iniidae | NCBI | RJWO010011616.1 |
| 15 | *Lipotes vexillifer* | Lipotidae | NCBI | NW\_006772141.1 |
| 16 | *Mesoplodon bidens* | Ziphiidae | NCBI | PVJJ010000221.1 |
| 17 | *Ziphius cavirostris* | NCBI | RJWS010151175.1 |
| 18 | *Kogia breviceps* | Kogiidae | NCBI | RJWL010050176.1 |
| 19 | *Physeter macrocephalus* | Physeteridae | NCBI | NC\_041224.1 |
| 20 | *Balaenoptera edeni* | Balaenopteridae | DNA ZOO | HiC\_scaffold\_20 |
| 21 | *Balaenoptera musculus* | NCBI | CM020942.1 |
| 22 | *Balaenoptera physalus* | NCBI | SGJD01002715.1 |
| 23 | *Megaptera novaeangliae* | NCBI | RYZJ01000659.1 |
| 24 | *Eschrichtius robustus* | Eschrichtiidae | NCBI | NIPP01005531.1 |
| 25 | *Balaenoptera acutorostrata scammoni* | Balaenopteridae | NCBI | NW\_006726576.1 |
| 26 | *Balaenoptera bonaerensis* | NCBI | DF437389.1 |
| 27 | *Eubalaena japonica* | Balaenidae | DNA ZOO | RJWP010000613.1 |
| 28 | *Eubalaena glacialis* | DNA ZOO | HiC\_scaffold\_13 |
| 29 | *Balaena mysticetus* | bowhead-whale.org | scaffold\_1635 |
| 30 | *Hippopotamus amphibius* | Hippopotamidae | NCBI | PVJP02000056.1 |

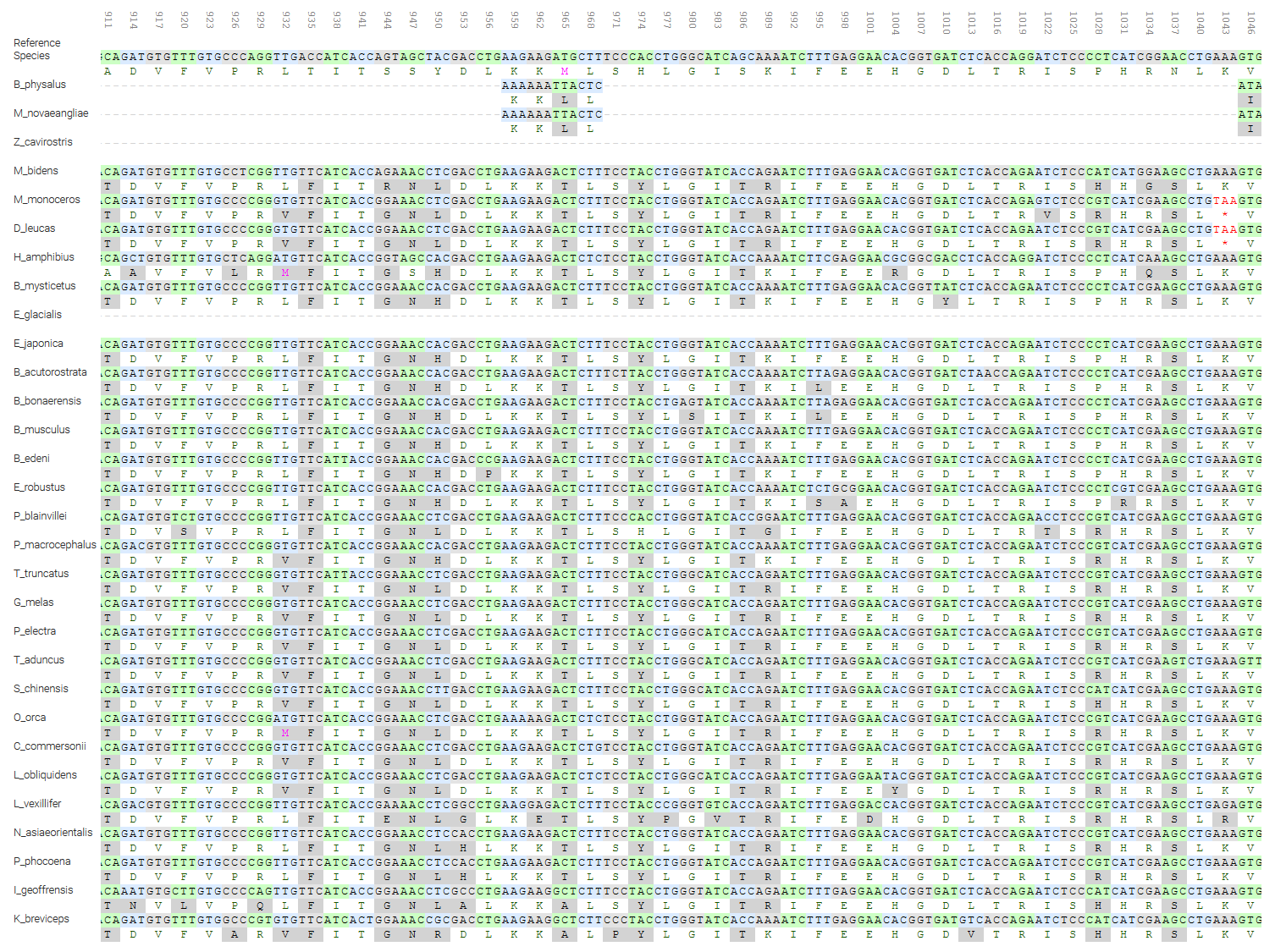
**Supplementary Table 2:** Analyzed genomic sequences regarding the Pseudo*Checker*’s coding status assessment of *SLURP1* in cetaceans and *Hippopotamus amphibius*.

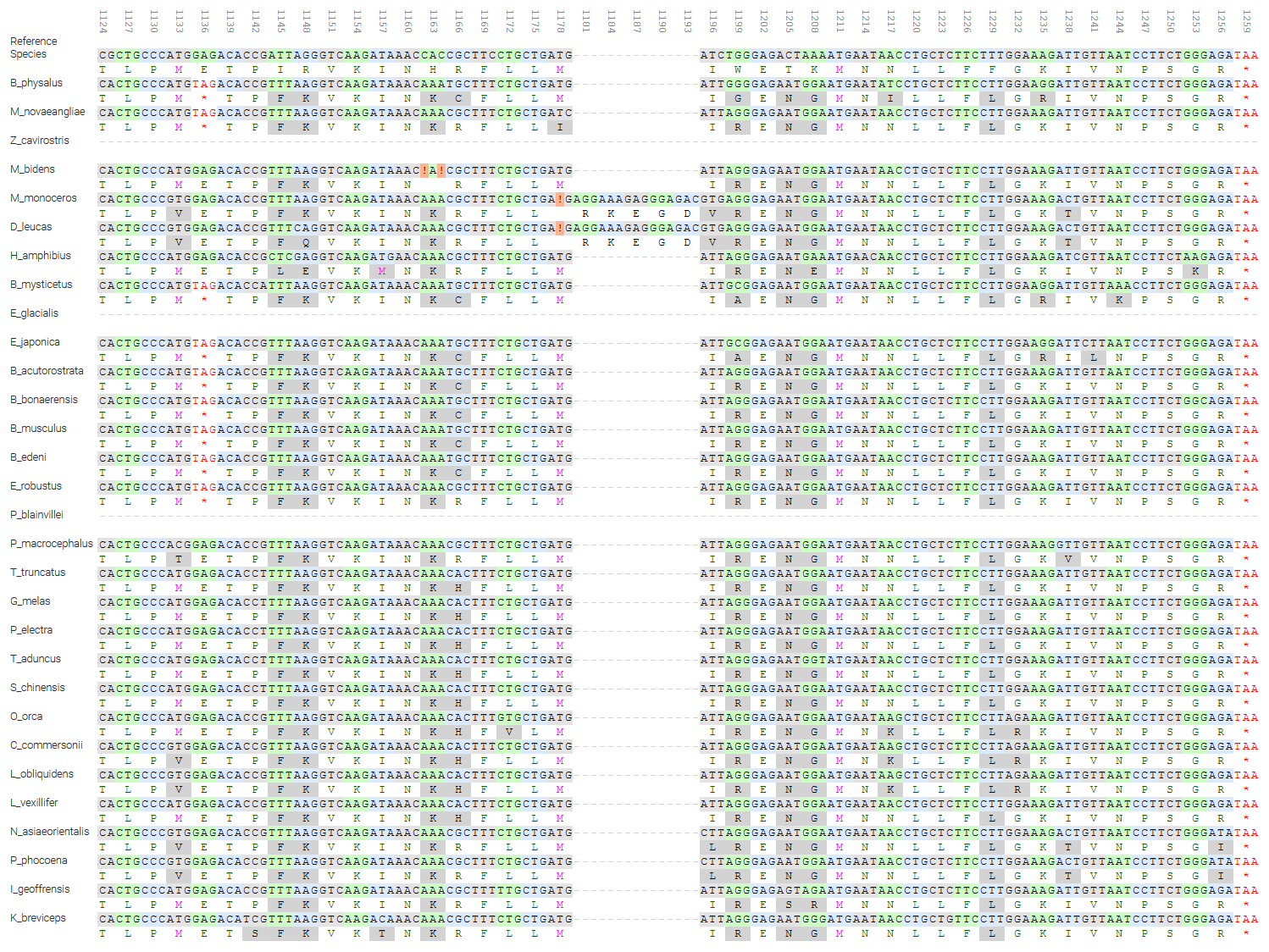
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| # | Species | Family | Database | Genomic Sequence ID |
| 1 | *Tursiops aduncus* | Delphinidae | NCBI | NCQN01016125.1 |
| 2 | *Tursiops truncatus* | NCBI | NC\_047050.1 |
| 3 | *Sousa chinensis* | NCBI | RWJT01019188.1 |
| 4 | *Globicephala melas* | NCBI | NW\_022134900.1 |
| 5 | *Peponocephala electra* | DNA ZOO | HiC\_scaffold\_12 |
| 6 | *Cephalorhynchus commersonii* | DNA ZOO | HiC\_scaffold\_20 |
| 7 | *Lagenorhynchus obliquidens* | NCBI | NW\_020837985.1 |
| 8 | *Orcinus orca* | NCBI | NW\_004438426.1 |
| 9 | *Neophocaena asiaeorientalis asiaeorientalis* | Phocoenidae | NCBI | NW\_020172779.1 |
| 10 | *Phocoena phocoena* | NCBI | PKGA01139387.1 |
| 11 | *Monodon monoceros* | Monodontidae | NCBI | NW\_021703782.1 |
| 12 | *Delphinapterus leucas* | NCBI | NW\_022098049.1 |
| 13 | *Pontoporia blainvillei* | Pontoporiidae | NCBI | PGGH01252053.1 |
| 14 | *Inia geoffrensis* | Iniidae | NCBI | RJWO010003874.1 |
| 15 | *Lipotes vexillifer* | Lipotidae | NCBI | NW\_006791731.1 |
| 16 | *Mesoplodon bidens* | Ziphiidae | NCBI | PVJJ010033365.1 |
| 17 | *Ziphius cavirostris* | NCBI | RJWS010002737.1 |
| 18 | *Kogia breviceps* | Kogiidae | NCBI | -- |
| 19 | *Physeter macrocephalus* | Physeteridae | NCBI | NW\_021146804.1 |
| 20 | *Balaenoptera edeni* | Balaenopteridae | DNA ZOO | HiC\_scaffold\_10 |
| 21 | *Balaenoptera musculus* | NCBI | CM020957.1 |
| 22 | *Balaenoptera physalus* | NCBI | SGJD01005450.1 |
| 23 | *Megaptera novaeangliae* | NCBI | RYZJ01000962.1 |
| 24 | *Eschrichtius robustus* | Eschrichtiidae | NCBI | NIPP01000017.1 |
| 25 | *Balaenoptera acutorostrata scammoni* | Balaenopteridae | NCBI | NW\_006725399.1 |
| 26 | *Balaenoptera bonaerensis* | NCBI | DF402240.1 |
| 27 | *Eubalaena japonica* | Balaenidae | DNA ZOO | RJWP010039703.1 |
| 28 | *Eubalaena glacialis* | DNA ZOO | HiC\_scaffold\_18 |
| 29 | *Balaena mysticetus* | bowhead-whale.org | scaffold\_2155 |
| 30 | *Hippopotamus amphibius* | Hippopotamidae | NCBI | PVJP02910145.1 |

**Supplementary Figure 1:** Pseudo*Checker*’s analysis screenshot: demonstration of a premature in-frame stop codon (alignment pos. 151) arising in exon 1 of *E. glacialis* and *E. japonica*’s *SERPINA12* gene.

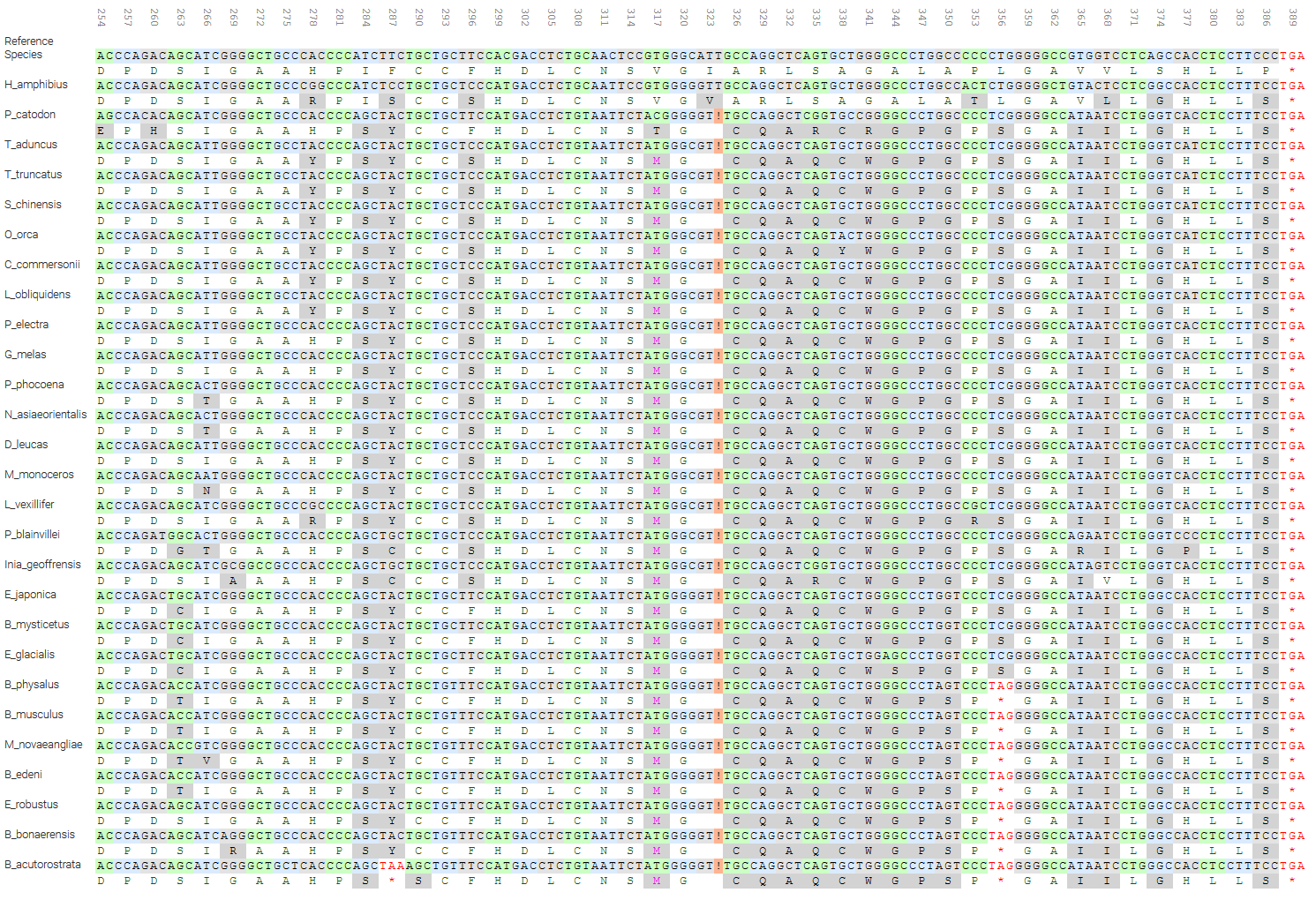


**Supplementary Figure 2:** Pseudo*Checker*’s analysis screenshot: Demonstration of a single-nucleotide deletion (alignment pos. 525) in exon 1 of *M. novaengliae*, *B. mysticetus, E. glacialis, E. japonica, B. edeni* and *E. robustus SERPINA12* gene, also including the demonstration of an in-frame premature stop codon (alignment pos. 631) arising in the same exon and gene of *B.* *acutorostrata scammoni*.

**Supplementary Figure 3:** Pseudo*Checker*’s analysis screenshot: Demonstration of a premature in-frame stop codon (alignment pos. 1042) arising in exon 3 of *M. monoceros* and *D. leucas*’s *SERPINA12* gene.



**Supplementary Figure 4:** Pseudo*Checker*’s analysis screenshot: Demonstration of (a): an in-frame premature stop codon arising (alignment pos. 1135) in exon 4 of *B. physalus, M. novaeangliae, B. mysticetus, E. japonica, B. acutorostrata scammoni, B. bonaerensis, B. musculus, B. edeni* and *E. robustus’s SERPINA12* gene; (b) two single-nucleotide deletions (alignment pos. 1162 and 1164) in exon 4 of *M. bidens’s* *SERPINA12* gene; (c) a single-nucleotide deletion (alignment pos. 1170) in exon 4 of *M. monoceros* and *D. leucas’s SERPINA12* gene.



**Supplementary Figure 5:** Pseudo*Checker*’s analysis screenshot: Demonstration of (a): an in-frame premature stop codon arising (alignment pos. 206) in exon 3 of *B. acutorostrata scammoni’s SLURP1* gene; (b) one single-nucleotide deletion (alignment pos. 324) in exon 3 of all represented species’ *SLURP1* gene, except for *H. amphibius*; (c) an in-frame premature stop codon arising (alignment pos. 355) in exon 3 of *B. physalus, B. musculus, M. novaengliae, B. edeni, E. robustus, B. bonaerensis* and *B. acutorostrata scammoni’s SLURP1* gene.