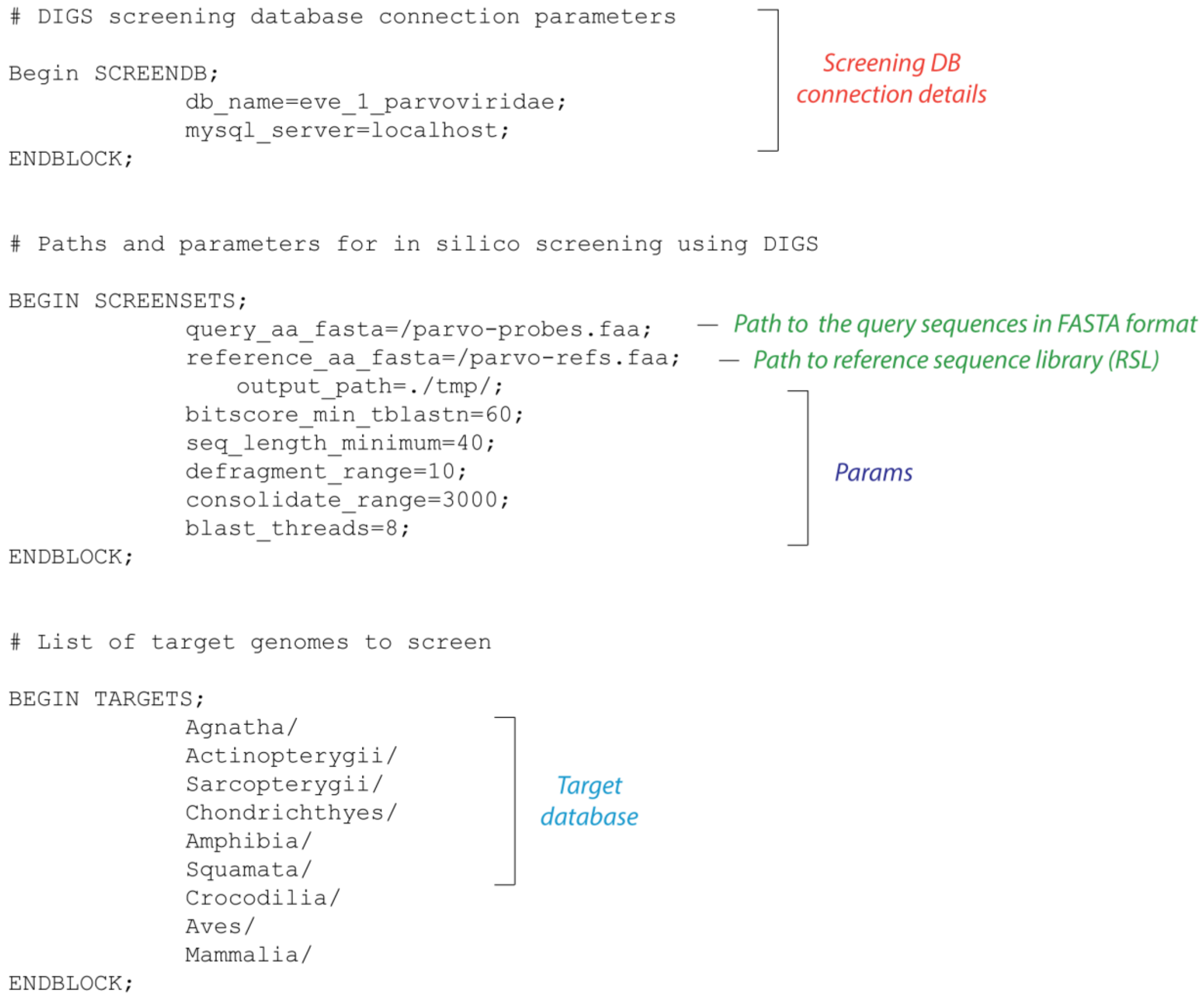

```
# DIGS screening database connection parameters

Begin SCREENDB;
            db_name=eve_1_parvoviridae;
            mysql_server=localhost;
ENDBLOCK;


# Paths and parameters for in silico screening using DIGS

BEGIN SCREENSETS;
            query_aa_fasta=/parvo-probes.faa;
            reference_aa_fasta=/parvo-refs.faa;
               output_path=./tmp/;
            bitscore_min_tblastn=60;
            seq_length_minimum=40;
            defragment_range=10;
            consolidate_range=3000;
            blast_threads=8;
ENDBLOCK;


# List of target genomes to screen

BEGIN TARGETS;
            Agnatha/
            Actinopterygii/
            Sarcopterygii/
            Chondrichthyes/
            Amphibia/
            Squamata/
            Crocodilia/
            Aves/
            Mammalia/
ENDBLOCK;
```

Figure S1. An annotated example of a DIGS tool control file.

The DIGS tool control file defines parameters for screening and the paths to reference sequence library (RSL), probe, and target database (TDb) files. Control files are structured as 'NEXUS' style blocks {Maddison, 1997 #86} delineated by BEGIN and ENDBLOCK tokens.

References:

Maddison, D.R., D.L. Swofford, and W.P. Maddison, *NEXUS: an extensible file format for systematic information.* Syst Biol, 1997. **46**(4): p. 590-621.