SWFSbase User Guide

SOCIETY FOR WILDLIFE FORENSIC SCIENCE

Table of Contents

Introduction

SWFSbase is a secure BLAST server developed for the wildlife forensic community. It centralises curated reference databases contributed by laboratories and links them to essential forensic metadata, while also incorporating selected publicly available reference data. SWFSbase is hosted on SequenceServer™ Cloud. More information about the platform is available on the SWFSbase webpage: https://www.wildlifeforensicscience.org/swfsbase/

This User Guide provides an overview of how to use SWFSbase, including the different interfaces of SWFSbase, how to run BLAST searches, and how to submit sequences.

If you have any questions or issues, please contact: SWFSbase@wildlifeforensicscience.org.

Accessing SWFSbase:

SWFSbase can only be used for wildlife forensic casework. It cannot be used for research unless the relevant sequences are 1) owned by you, 2) are publicly available, or 3) you have obtained permission from the data owner to use them for research.

To access SWFSbase you must:

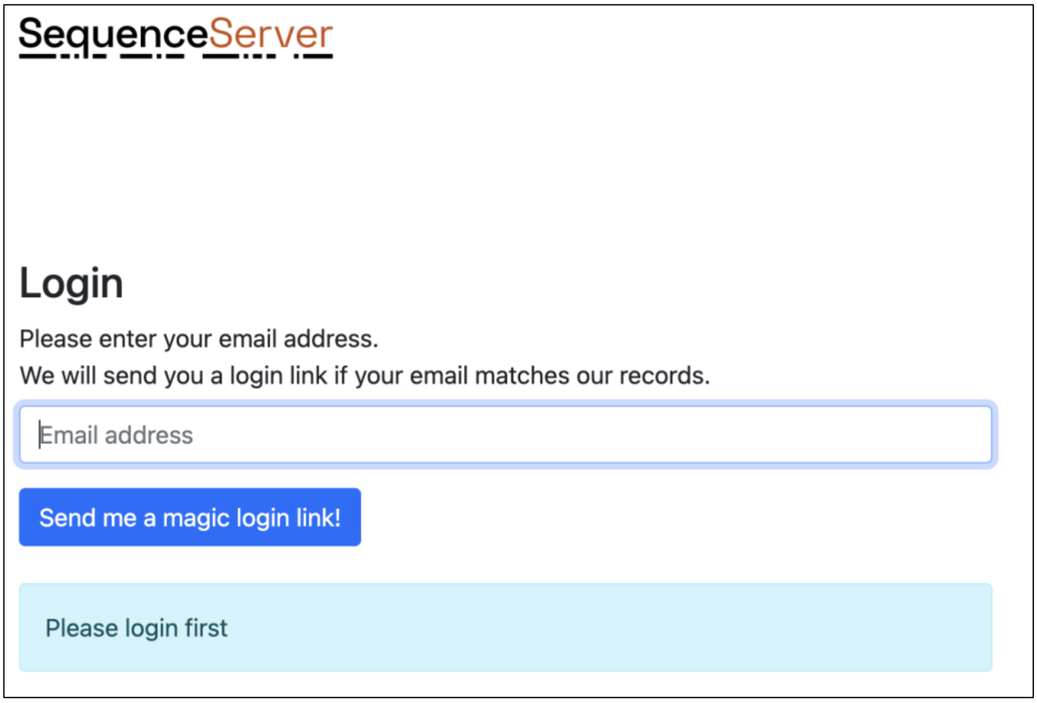
1. Be a member of the Society for Wildlife Forensic Science (SWFS).
2. Review and sign the SWFSbase Terms of Use.

The Terms of Use can be downloaded from the SWFSbase webpage. Signed agreements should be emailed to SWFSbase@wildlifeforensicscience.org. Please also include your full name, institution, and the email address you would like associated with your login.

Once these requirements have been completed, you will be added as a SWFSbase user. You can then access the platform using the following link: https://wildlifeforensics.sequenceserver.com

Logging in to SWFSbase

SWFSbase uses a secure login system (two-step verification). When accessing the platform using the link above, you will see the following page:



Login

Please enter your email address.
We will send you a login link if your email matches our records.

Please login first

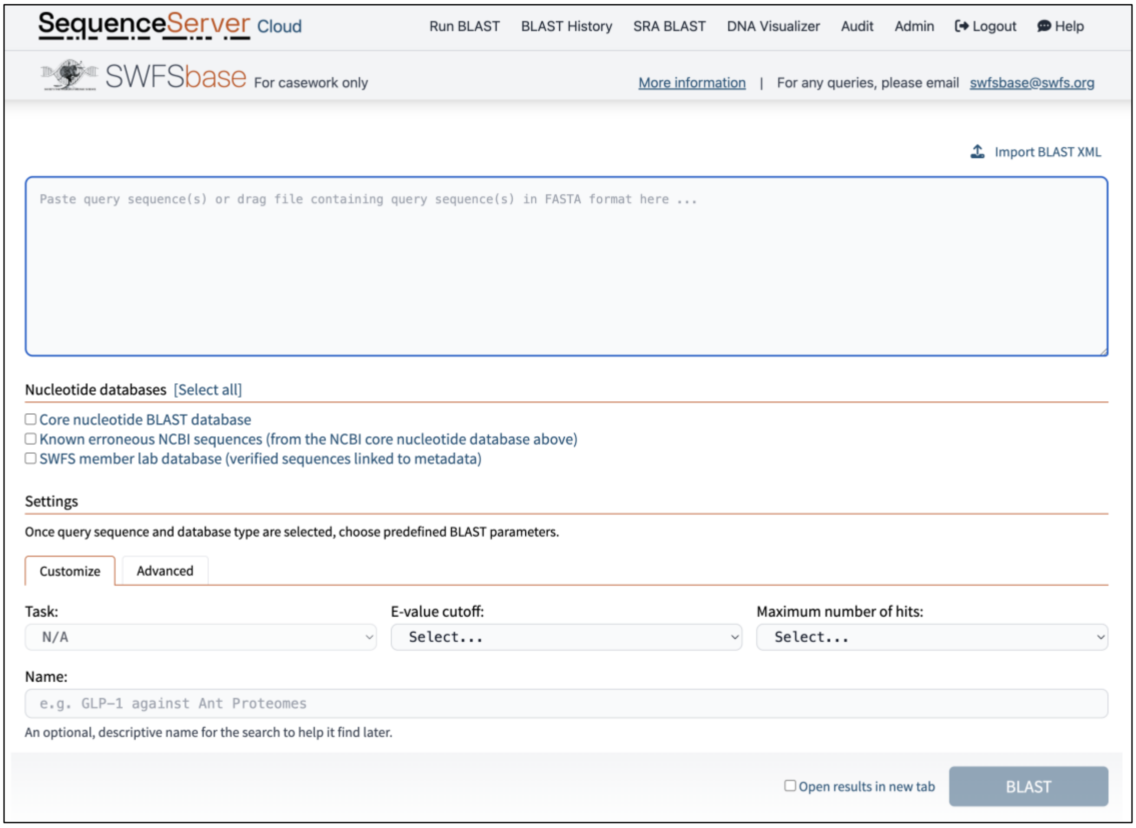
To log in:

1. Enter your registered email address.
2. You will receive an email containing a login link.
3. Click the link (or copy and paste it into your browser).

The login link remains valid for up to 30 days, or until you click Logout within SWFSbase. After this, you will need to repeat the login process. For security reasons, it is recommended that you log out after each session. Additional information about platform security can be found here.

Running a BLAST with SWFSbase

Once you are logged in you will see the following BLAST interface:



Paste query sequence(s) or drag file containing query sequence(s) in FASTA format here ...

Nucleotide databases [Select all]

☐ Core nucleotide BLAST database
☐ Known erroneous NCBI sequences (from the NCBI core nucleotide database above)
☐ SWFS member lab database (verified sequences linked to metadata)

Settings

Once query sequence and database type are selected, choose predefined BLAST parameters.

Task: N/A
E-value cutoff: Select...
Maximum number of hits: Select...

Name: e.g. GLP-1 against Ant Proteomes

An optional, descriptive name for the search to help it find later.

☐ Open results in new tab

This interface functions similarly to an NCBI BLAST search. You can paste your sequence/s directly into the query box or drag and drop a FASTA file. Multiple sequences can be submitted in a single search.

Included Databases

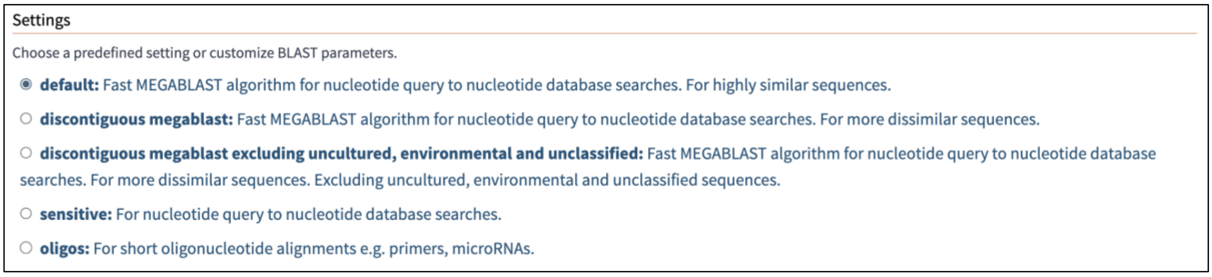
- ***SWFS member lab database (verified sequences linked to metadata):*** This is the key database within SWFSbase. It comprises vetted data contributed by multiple wildlife forensic laboratories, including data not available in other public repositories. Each sequence is linked to associated metadata to support assessment of specimen identification, sequence quality, and reliability (see image below). Laboratories may contribute sequences by following the submission instructions provided below.
- ***Core nucleotide BLAST database:*** This is the entire NCBI core nucleotide database ("core_nt"), updated following NCBI's frequency (typically daily to fortnightly). These sequences are not curated within SWFSbase and are not linked to forensic metadata. The database is provided for convenience, allowing users to conduct searches without leaving the platform.
- ***Known erroneous NCBI sequences (from the NCBI core nucleotide database above):*** This is a small database that contains sequences from the NCBI core nucleotide database that have previously been identified as potentially erroneous. After conducting a search against the standard NCBI database, users may wish to query this database to determine whether top matches correspond to problematic records.

Search Strategies

A useful strategy is to first search the ‘*SWFS member lab database’*, which contains highly reliable and transparent reference data. You may then wish to search the ‘*NCBI nucleotide database*’, which contains a much broader set of sequences. Finally, if your top matches come from NCBI, you can search the *‘Known erroneous NCBI sequences’* database to determine whether any of those matches correspond to previously identified problematic records.

Note that the larger the database used, the longer the BLAST search will take. The ‘*NCBI core nucleotide database*’ is by far the largest, and searches will therefore typically run slower than the other two databases. However, repeated searches against this database may run faster after the first search due to caching.

Once you enter a sequence, there are several BLAST presets available:



Settings

Choose a predefined setting or customize BLAST parameters.

- ● **default:** Fast MEGABLAST algorithm for nucleotide query to nucleotide database searches. For highly similar sequences.
- ○ **discontiguous megablast:** Fast MEGABLAST algorithm for nucleotide query to nucleotide database searches. For more dissimilar sequences.
- ○ **discontiguous megablast excluding uncultured, environmental and unclassified:** Fast MEGABLAST algorithm for nucleotide query to nucleotide database searches. For more dissimilar sequences. Excluding uncultured, environmental and unclassified sequences.
- ○ **sensitive:** For nucleotide query to nucleotide database searches.
- ○ **oligos:** For short oligonucleotide alignments e.g. primers, microRNAs.

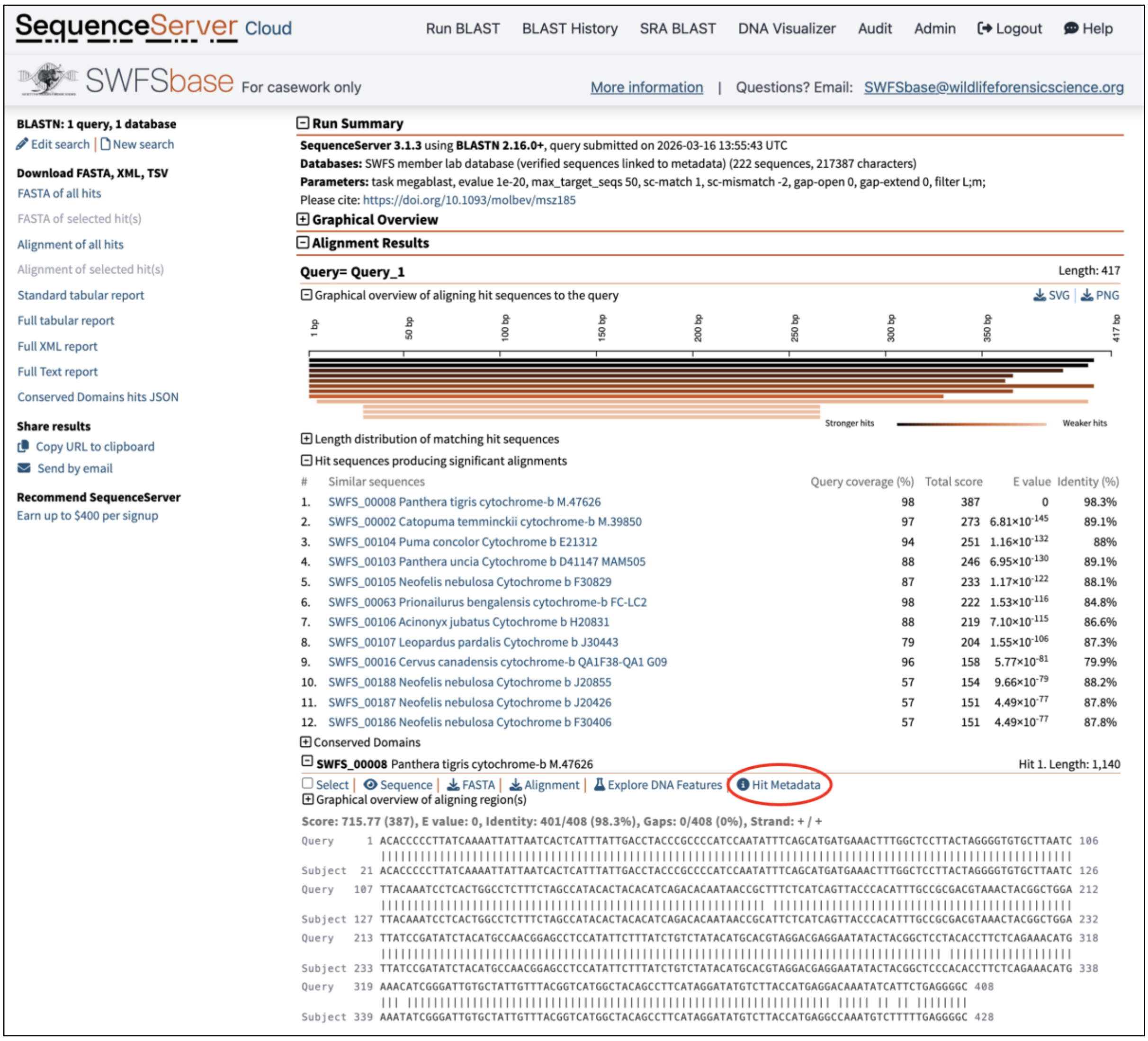
The ***‘default’*** settings are appropriate for most searches, and are designed to quickly find highly similar sequences. If BLASTing against the ‘*NCBI nucleotide database*’, it is useful to use the ***‘discontiguous megablast excluding uncultured, environmental and unclassified’*** setting; this allows mismatches in the initial alignment seed, making the search more sensitive to distant sequence similarity, while excluding many low-quality or irrelevant sequences (searches may run slightly slower than the default preset).

Users may also click the ‘Customize’ and ‘Advanced’ settings to manually adjust BLAST parameters and filters.

Viewing results

The BLAST results interface in SWFSbase is largely similar to NCBI BLAST. Several result fields are displayed by default; some fields may be irrelevant for your analysis and can be folded away using the minus signs. Your browser cookies should remember which fields you have folded, so they will remain folded in subsequent searches.

Below is an example of a BLAST result (the query is a tiger cytochrome-b sequence, searched against the ‘*SWFS member lab database*’), with some result fields already folded away:



SequenceServer 3.1.3 using **BLASTN 2.16.0+**, query submitted on 2026-03-16 13:55:43 UTC
Databases: SWFS member lab database (verified sequences linked to metadata) (222 sequences, 217387 characters)
Parameters: task megablast, evalue 1e-20, max_target_seqs 50, sc-match 1, sc-mismatch -2, gap-open 0, gap-extend 0, filter L;m;
Please cite: https://doi.org/10.1093/molbev/msz185

Query= Query_1 Length: 417

Graphical overview of aligning hit sequences to the query

#	Similar sequences	Query coverage (%)	Total score	E value	Identity (%)
1.	SWFS_00008 Panthera tigris cytochrome-b M.47626	98	387	0	98.3%
2.	SWFS_00002 Catopuma temminckii cytochrome-b M.39850	97	273	6.81×10^{-145}	89.1%
3.	SWFS_00104 Puma concolor Cytochrome b E21312	94	251	1.16×10^{-132}	88%
4.	SWFS_00103 Panthera uncia Cytochrome b D41147 MAM505	88	246	6.95×10^{-130}	89.1%
5.	SWFS_00105 Neofelis nebulosa Cytochrome b F30829	87	233	1.17×10^{-122}	88.1%
6.	SWFS_00063 Prionailurus bengalensis cytochrome-b FC-LC2	98	222	1.53×10^{-116}	84.8%
7.	SWFS_00106 Acinonyx jubatus Cytochrome b H20831	88	219	7.10×10^{-115}	86.6%
8.	SWFS_00107 Leopardus pardalis Cytochrome b J30443	79	204	1.55×10^{-106}	87.3%
9.	SWFS_00016 Cervus canadensis cytochrome-b QA1F38-QA1 G09	96	158	5.77×10^{-81}	79.9%
10.	SWFS_00188 Neofelis nebulosa Cytochrome b J20855	57	154	9.66×10^{-79}	88.2%
11.	SWFS_00187 Neofelis nebulosa Cytochrome b J20426	57	151	4.49×10^{-77}	87.8%
12.	SWFS_00186 Neofelis nebulosa Cytochrome b F30406	57	151	4.49×10^{-77}	87.8%

SWFS_00008 Panthera tigris cytochrome-b M.47626 Hit 1. Length: 1,140

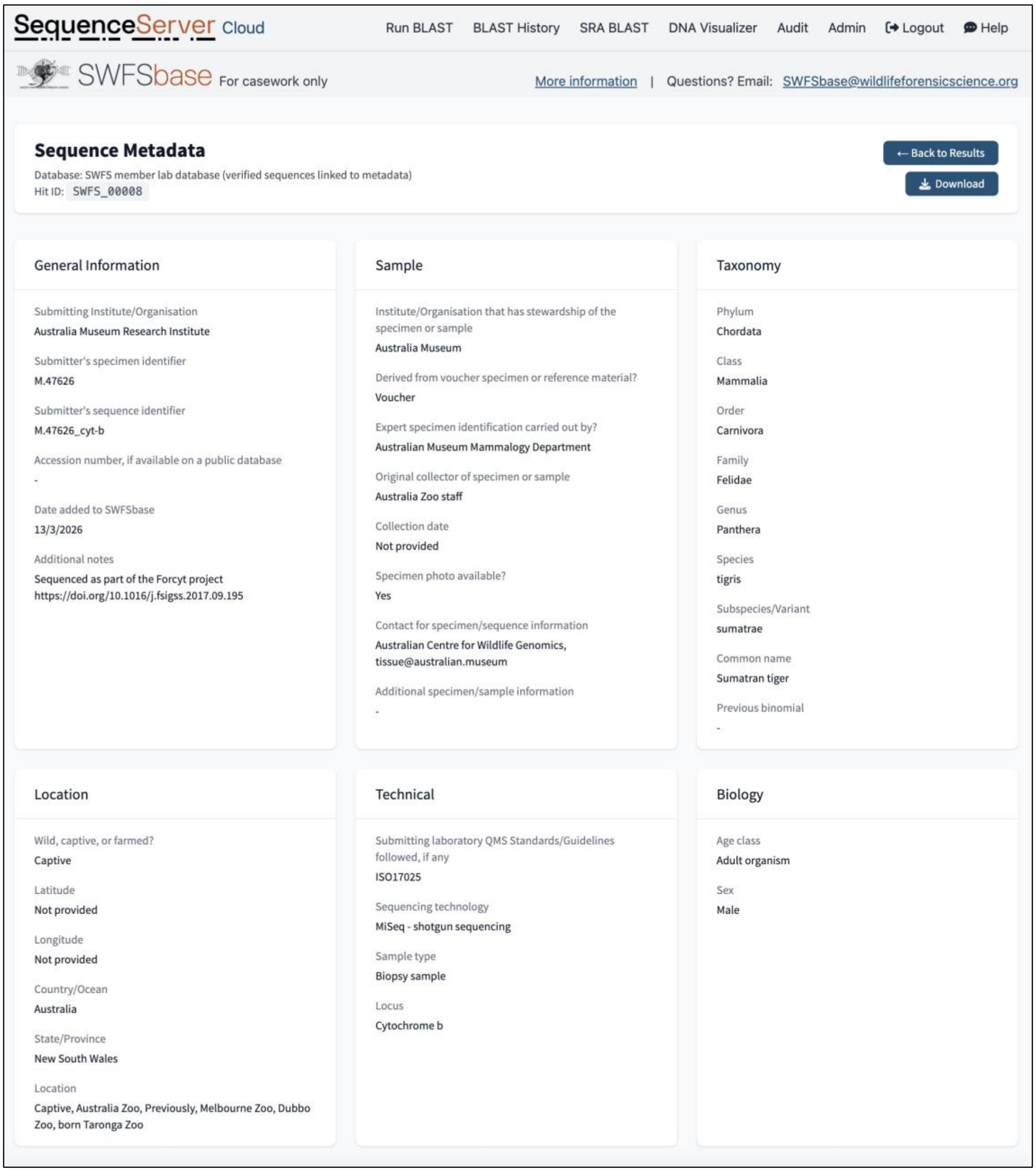
Score: 715.77 (387), E value: 0, Identity: 401/408 (98.3%), Gaps: 0/408 (0%), Strand: + / +

```
Query     1 ACACCCCCTTATCAAAATTATTAATCACTCATTTATTGACCTACCCGCCCCATCCAATATTTCAGCATGATGAAACTTTGGCTCCTTACTAGGGGTGTGCTTAATC 106
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Subject  21 ACACCCCCTTATCAAAATTATTAATCACTCATTTATTGACCTACCCGCCCCATCCAATATTTCAGCATGATGAAACTTTGGCTCCTTACTAGGGGTGTGCTTAATC 126
Query   107 TTACAAATCCTCACTGGCCTCTTTCTAGCCATACACTACACATCAGACACAATAACCGCTTTCTCATCAGTTACCCACATTTGCCGCGACGTAAACTACGGCTGGA 212
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Subject 127 TTACAAATCCTCACTGGCCTCTTTCTAGCCATACACTACACATCAGACACAATAACCGCATTCTCATCAGTTACCCACATTTGCCGCGACGTAAACTACGGCTGGA 232
Query   213 TTATCCGATATCTACATGCCAACGGAGCCTCCATATTCTTTATCTGTCTATACATGCACGTAGGACGAGGAATATACTACGGCTCCTACACCTTCTCAGAAACATG 318
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Subject 233 TTATCCGATATCTACATGCCAACGGAGCCTCCATATTCTTTATCTGTCTATACATGCACGTAGGACGAGGAATATACTACGGCTCCCACACCTTCTCAGAAACATG 338
Query   319 AAACATCGGGATTGTGCTATTGTTTACGGTCATGGCTACAGCCTTCATAGGATATGTCTTACCATGAGGACAAATATCATTCTGAGGGGC 408
            ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| || || ||||||||
Subject 339 AAATATCGGGATTGTGCTATTGTTTACGGTCATGGCTACAGCCTTCATAGGATATGTCTTACCATGAGGCCAAATGTCTTTTTGAGGGGC 428
```

The bottom of the screenshot displays the sequence alignment for the first sequence hit; scrolling down will show additional hits.

Metadata

If a sequence has metadata (as do all sequences in the '*SWFS member lab database*'), it will display a "Hit Metadata" link (circled in red above). Clicking this link opens the metadata interface, where you can view detailed information about the sequence. Below is the metadata for the top hit from the example search:



Sequence Metadata

Database: SWFS member lab database (verified sequences linked to metadata)
Hit ID: `SWFS_00008`

General Information

Submitting Institute/Organisation
Australia Museum Research Institute

Submitter's specimen identifier
M.47626

Submitter's sequence identifier
M.47626_cyt-b

Accession number, if available on a public database
-

Date added to SWFSbase
13/3/2026

Additional notes
Sequenced as part of the Forcyt project https://doi.org/10.1016/j.fsigss.2017.09.195

Sample

Institute/Organisation that has stewardship of the specimen or sample
Australia Museum

Derived from voucher specimen or reference material?
Voucher

Expert specimen identification carried out by?
Australian Museum Mammalogy Department

Original collector of specimen or sample
Australia Zoo staff

Collection date
Not provided

Specimen photo available?
Yes

Contact for specimen/sequence information
Australian Centre for Wildlife Genomics, tissue@australian.museum

Additional specimen/sample information
-

Taxonomy

Phylum
Chordata

Class
Mammalia

Order
Carnivora

Family
Felidae

Genus
Panthera

Species
tigris

Subspecies/Variant
sumatrae

Common name
Sumatran tiger

Previous binomial
-

Location

Wild, captive, or farmed?
Captive

Latitude
Not provided

Longitude
Not provided

Country/Ocean
Australia

State/Province
New South Wales

Location
Captive, Australia Zoo, Previously, Melbourne Zoo, Dubbo Zoo, born Taronga Zoo

Technical

Submitting laboratory QMS Standards/Guidelines followed, if any
ISO17025

Sequencing technology
MiSeq - shotgun sequencing

Sample type
Biopsy sample

Locus
Cytochrome b

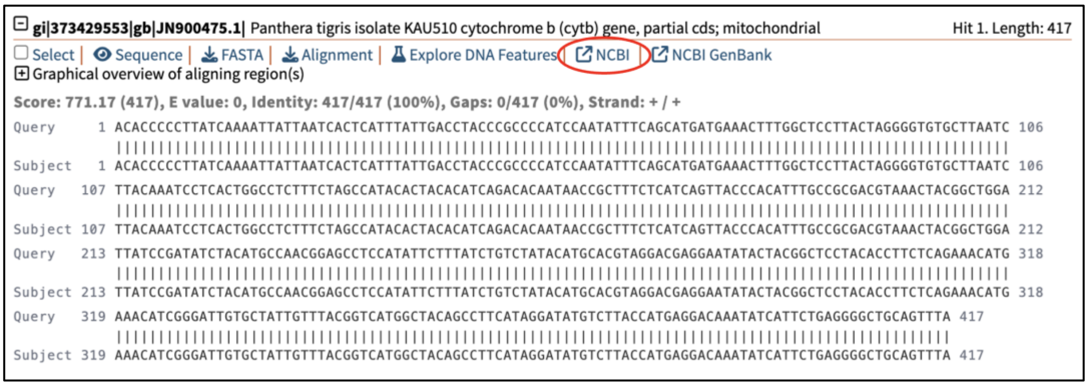
Biology

Age class
Adult organism

Sex
Male

These metadata are intended to provide information about the reliability and provenance of the sequence hit, supporting interpretation of your results.

Sequences from the NCBI core nucleotide database are not linked to metadata within SWFSbase. To access additional information for these hits, you can click the NCBI link provided to view the sequence details on the NCBI website:



⊟ **gi|373429553|gb|JN900475.1|** Panthera tigris isolate KAU510 cytochrome b (cytb) gene, partial cds; mitochondrial Hit 1. Length: 417

Score: 771.17 (417), E value: 0, Identity: 417/417 (100%), Gaps: 0/417 (0%), Strand: + / +

```
Query     1 ACACCCCCTTATCAAAATTATTAATCACTCATTTATTGACCTACCCGCCCCATCCAATATTTCAGCATGATGAAACTTTGGCTCCTTACTAGGGGTGTGCTTAATC 106
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Subject   1 ACACCCCCTTATCAAAATTATTAATCACTCATTTATTGACCTACCCGCCCCATCCAATATTTCAGCATGATGAAACTTTGGCTCCTTACTAGGGGTGTGCTTAATC 106
Query   107 TTACAAATCCTCACTGGCCTCTTTCTAGCCATACACTACACATCAGACACAATAACCGCTTTCTCATCAGTTACCCACATTTGCCGCGACGTAAACTACGGCTGGA 212
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Subject 107 TTACAAATCCTCACTGGCCTCTTTCTAGCCATACACTACACATCAGACACAATAACCGCTTTCTCATCAGTTACCCACATTTGCCGCGACGTAAACTACGGCTGGA 212
Query   213 TTATCCGATATCTACATGCCAACGGAGCCTCCATATTCTTTATCTGTCTATACATGCACGTAGGACGAGGAATATACTACGGCTCCTACACCTTCTCAGAAACATG 318
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Subject 213 TTATCCGATATCTACATGCCAACGGAGCCTCCATATTCTTTATCTGTCTATACATGCACGTAGGACGAGGAATATACTACGGCTCCTACACCTTCTCAGAAACATG 318
Query   319 AAACATCGGGATTGTGCTATTGTTTACGGTCATGGCTACAGCCTTCATAGGATATGTCTTACCATGAGGACAAATATCATTCTGAGGGGCTGCAGTTTA 417
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Subject 319 AAACATCGGGATTGTGCTATTGTTTACGGTCATGGCTACAGCCTTCATAGGATATGTCTTACCATGAGGACAAATATCATTCTGAGGGGCTGCAGTTTA 417
```

Submitting sequences to SWFSbase

Laboratories that contribute data to SWFSbase play a vital role in enhancing the platform and supporting its mission.

To contribute to the '*Known erroneous NCBI sequence*' database, please report any identified erroneous NCBI sequences to SWFSbase@wildlifeforensicscience.org.

To contribute your laboratory's sequence data and associated metadata suitable for forensic analysis, please follow the instructions below.

Each submission consists of two files: (1) a FASTA file containing sequences (.fasta, .fa, .fas, or .fsa), and (2) a metadata file prepared in Excel using the metadata template available on the SWFSbase webpage.

Metadata entry

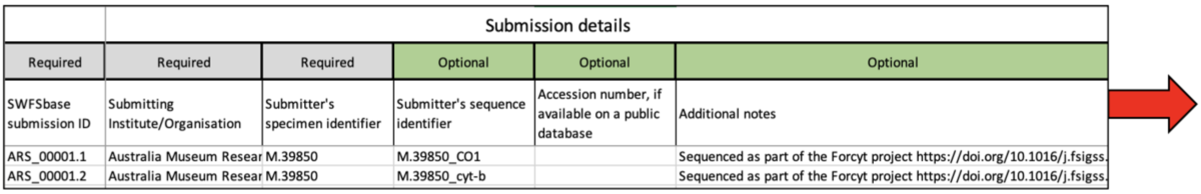
Information about each metadata field is provided on the first tab of the template. Please contact the SWFSbase team if you are unsure about any field.

Linking metadata to your FASTA

A key field in the metadata is '***SWFSbase submission ID***', which links each metadata row to its corresponding sequence. Each row of metadata in the Excel sheet must have a unique 'SWFSbase submission ID' that matches the ***sequence ID*** in the FASTA file.

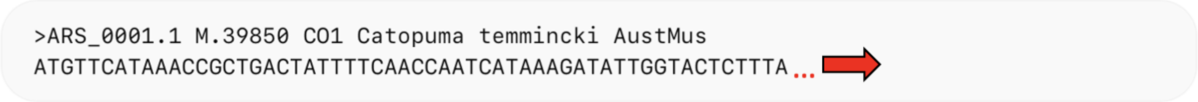
Note: *In a FASTA file, the sequence ID is the text immediately following the* **>** *symbol in the header, up to the first space. Anything after the first space is ignored.*

For example, the top row of this metadata excel template screenshot has a 'SWFSbase submission ID' `ARS_0001.1`:



	Submission details				
Required	Required	Required	Optional	Optional	Optional
SWFSbase submission ID	Submitting Institute/Organisation	Submitter's specimen identifier	Submitter's sequence identifier	Accession number, if available on a public database	Additional notes
ARS_00001.1	Australia Museum Resear	M.39850	M.39850_CO1		Sequenced as part of the Forcyt project https://doi.org/10.1016/j.fsigss.
ARS_00001.2	Australia Museum Resear	M.39850	M.39850_cyt-b		Sequenced as part of the Forcyt project https://doi.org/10.1016/j.fsigss.

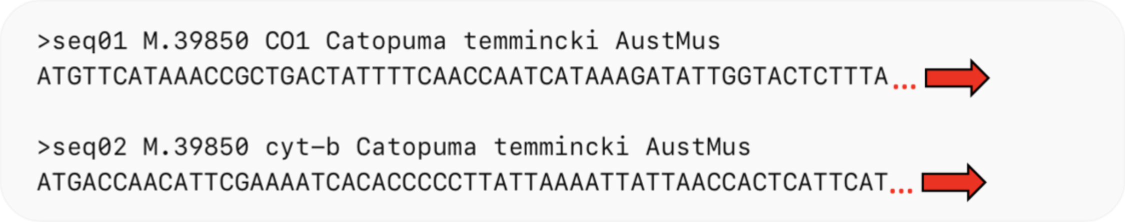
This should match the associated sequence in the FASTA file (the full sequence isn't displayed here):



```
>ARS_0001.1 M.39850 CO1 Catopuma temmincki AustMus
ATGTTCATAAACCGCTGACTATTTTCAACCAATCATAAAGATATTGGTACTCTTTA...
```

Note, any labelling system can be used for your 'SWFSbase submission IDs', as long as they are unique within that submission. You may use existing sequence IDs from your database (in this case, the 'SWFSbase submission ID' might be the same as the 'Submitter's specimen identifier' and/or 'Submitter's sequence identifier').

Alternatively, you can assign sequential IDs for your sequences. In the example below, the sequences are both from specimen M.39850, but for different loci, so `seq01`, and `seq02` have been used as the 'SWFSbase submission IDs' in both the FASTA and metadata files:



```
>seq01 M.39850 CO1 Catopuma temmincki AustMus
ATGTTCATAAACCGCTGACTATTTTCAACCAATCATAAAGATATTGGTACTCTTTA...

>seq02 M.39850 cyt-b Catopuma temmincki AustMus
ATGACCAACATTCGAAAATCACACCCCCTTATTAAAATTATTAACCACTCATTCAT...
```

Important:

- Ensure that all 'SWFSbase submission IDs' are unique and match exactly between the metadata and sequence files.
- Contact the SWFSbase team if you are unsure about formatting or any field in the template.

Once your sequences and metadata are prepared:

- Open the submission form here: https://docs.google.com/forms/d/e/1FAIpQLSfKaROuBi3QilkBdS3xsBclujdTqFhFb304JcMC5LAVPeFIWA/viewform
- Fill in the required fields.
- Attach your FASTA sequence file and metadata Excel file to the form.
- Submit the form.

The SWFSbase team will review each submission and contact you if there are any issues