

Genomics of farm animals/ Sequencing demonstration

Lecture

Modul no. 1: Animal Genetics

Aleš Knoll

Mendel University in Brno, Faculty of AgriScience



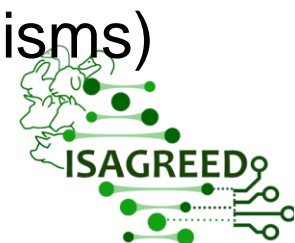
Co-funded by
the European Union

DNA sequencing

= determining the sequence of DNA nucleotides

- a) chemical method (formerly)
- b) ddNTP method, cyclic reaction (asymmetric PCR), today prevails, also Sanger sequencing (so-called 1st generation)
- c) pyrosequencing
- d) sequencing by hybridization
- e) NGS - sequencing of the new generation, modern large-format applications, parallel sequencing of fragments (2nd generation)
- f) single molecule sequencing (3rd generation)

- accurate identification of the polymorphic site
- automatic sequencers (also for direct detection of polymorphisms)



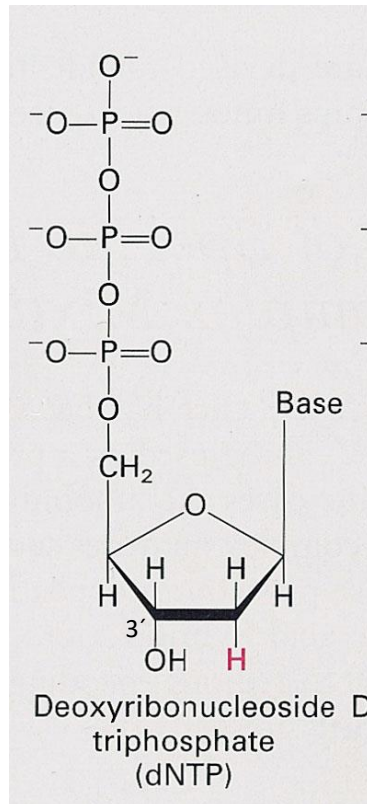
Co-funded by
the European Union

Sanger sequencing

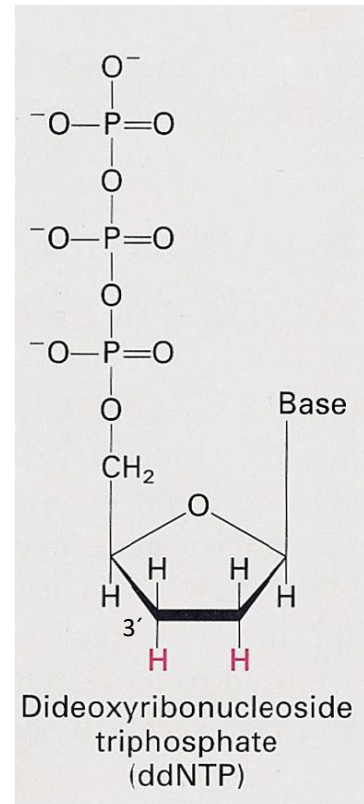
chain termination method: dideoxynucleoside triphosphates (ddNTPs)
they are labeled with 4 fluorescent colors according to the base



Frederick Sanger



standard nucleotide (dNTP)
enables the binding of another
nucleotide to the 3' end



modified nucleotide (ddNTP)
terminates chain synthesis
after binding

Erasmus+ project 2021-1-SK01-KA220-HED-000032068



Co-funded by
the European Union

Overview of sequencing process

1. Sample preparation: usually PCR product

1.1. Sample cleanup

1.2. Determination of DNA template concentration

2. Sequencing reaction - cyclic

sequencing primer (1 only)!

polymerase

buffer (+Mg²⁺)

dNTPs

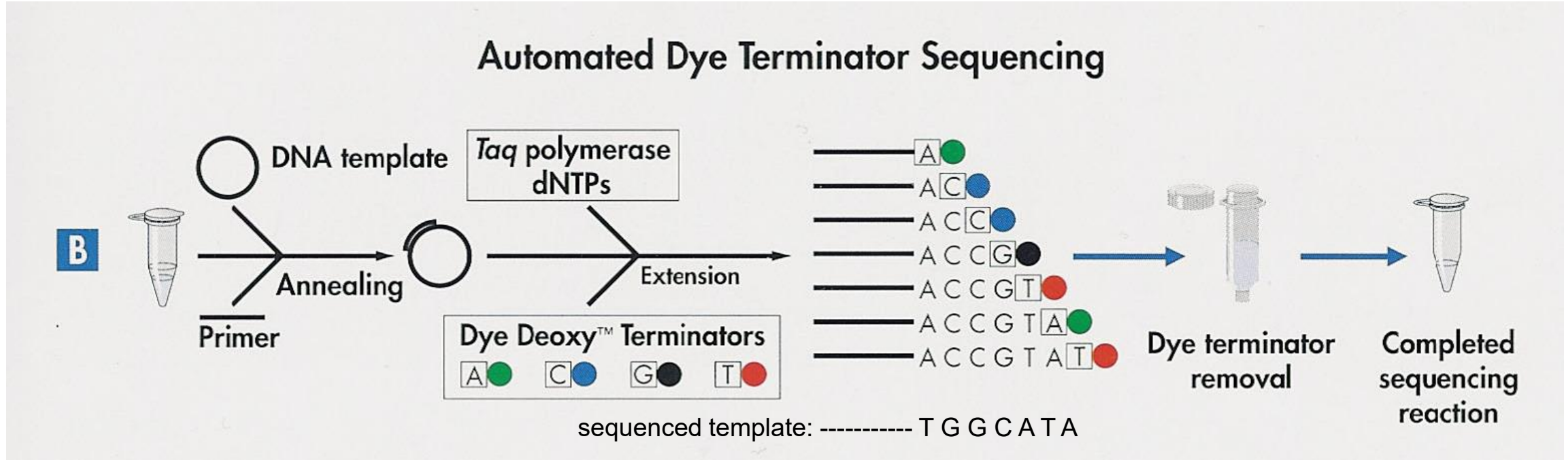
labelled ddNTP

3. Purification of the sequencing mixture – removal of free ddNTPs

4. Electrophoresis (capillary) on the sequencer

5. (Automatic) evaluation - determining the sequence

Four-color terminator sequencing

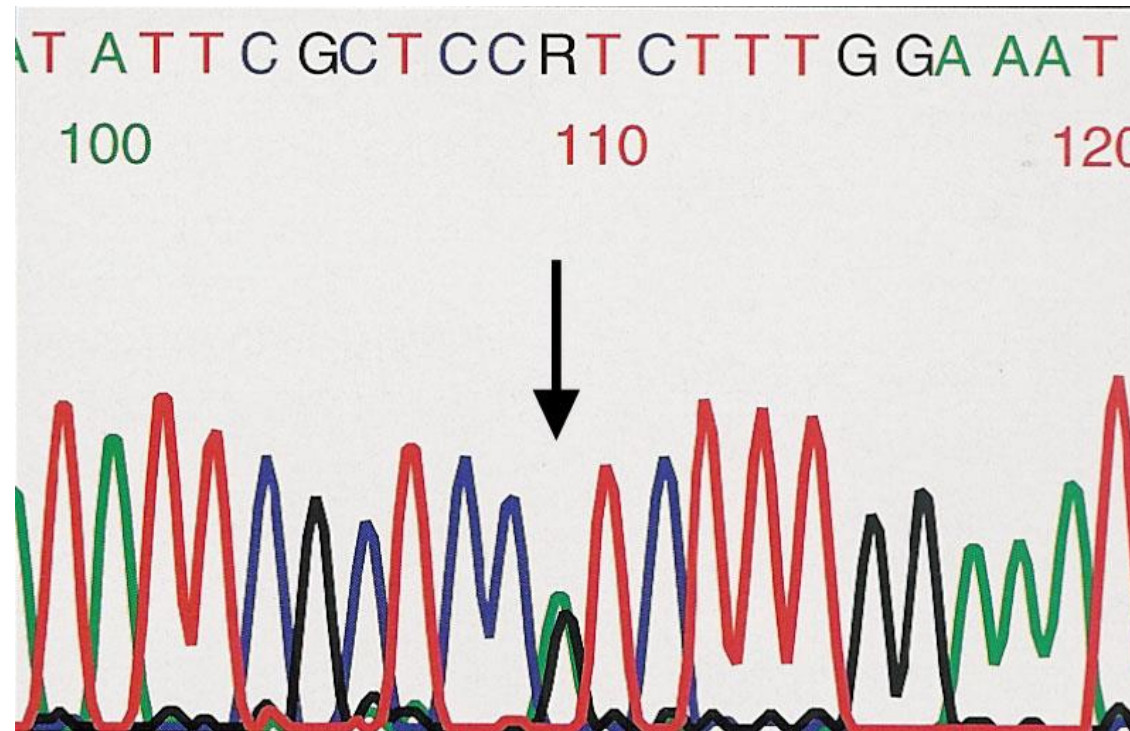


labeled terminators, everything takes place in 1 reaction, including electrophoresis, current variant

Sequencer

By capillary electrophoresis, single strands of different lengths are separated by size and the corresponding fluorescent color is read, which indicates the base present at that DNA site.

E.g., the 100 nucleotides long chain emits green light, i.e., there was an adenine at position 100 of the template



Advantages and disadvantages of Sanger sequencing

- long reads (up to 800 bp)
- high accuracy, reliability
- low overall price
- low performance, high price per base sequenced
- suitable for sequencing individual genes and detecting mutations in known sequences



Co-funded by
the European Union

New generation sequencers

so-called "New Generation Sequencing", NGS

second generation sequencing

- used for whole-genome sequencing
- they sequence in small pieces, but in a huge number of these fragments at the same time (massively parallel sequencing)
- enables a high capacity of sequencers and a significant reduction in the price per sequenced base
- e.g. Illumina; Life Technologies SOLiD 3 and more



Co-funded by
the European Union

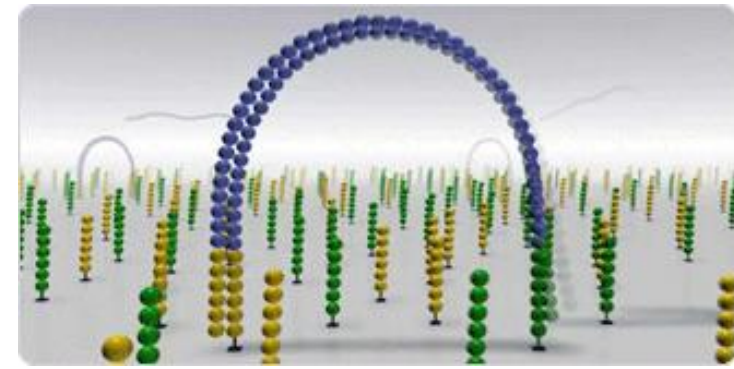
NGS Illumina

Illumina: NovSeq : aprox. 8 TB

NextSeq: up to 120 GB (formerly HiSeq 600GB)

MiSeq: up to 15 GB

MiniSeq: 7,5 GB



*sequencing by synthesis

Principle: bridge PCR, 4 SBS* fluorescence sequencing, read 150 to 300 bp, sequencing in clusters of the same sequence– the corresponding spot lights up in color according to NTP.



NextSeq

MiSeq



Co-funded by
the European Union

NGS summary

- the size of mammalian genomes is around 3 billion bp
- covers the entire genetic information of an individual
- allows finding new or identifying existing polymorphisms and mutations
- the problem is a large amount of data and its interpretation
- results stored in genomic databases (ncbi etc.)
- possibility of individual sequencing (personal genomics)



Co-funded by
the European Union

Third-generation sequencing

Single Molecule Technologies

- sequencing of individual molecules
- there are no amplification errors
- long reads (de novo sequencing) – up to 7kb!!! (3.5-8)
- a small amount of template
- high reading error rate so far (that's why the sequence must be read repeatedly)

example:

SMRT – Single Molecule Real Time Sequencing (fa. Pacific Bioscience)

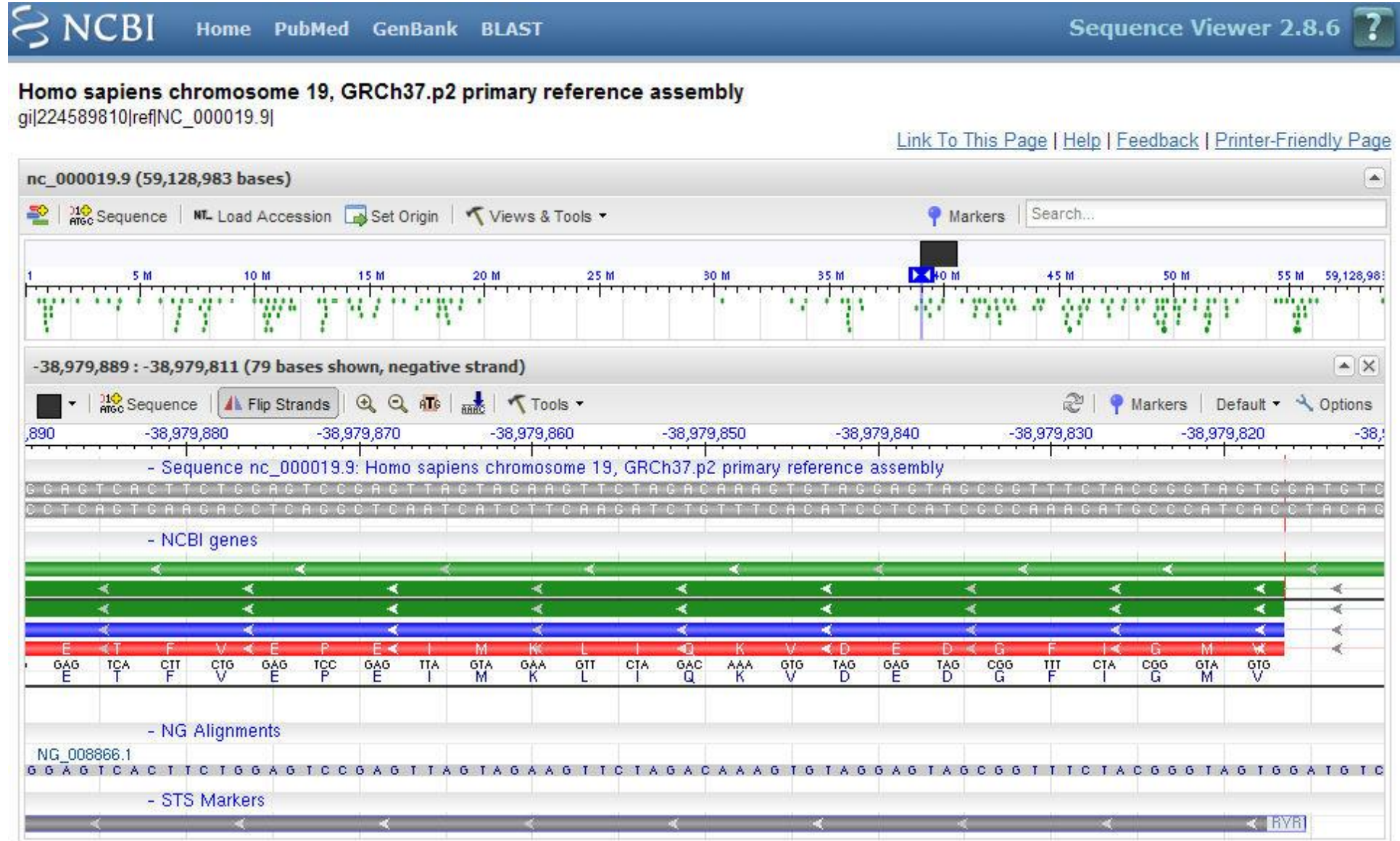
nanopore technology (Oxford Nanopore)



Co-funded by
the European Union

Physical map from genome sequencing

the NCBI Internet database allows viewing of genes down to the nucleotide sequence level



IS AGREED



Co-funded by
the European Union

Species identification using sequencing

Molecular taxonomy, **DNA barcoding**

makes it possible to identify a species where it is not possible with the classical method, e.g. in the case of insect larvae

a) mtDNA (COI)

b) nuclear genes

Sequencing of a specific fragment, comparison in Bold or Blast databases.

The advantage of mtDNA is a multicopy, i.e. greater amounts of DNA per amount of biol. material and higher stability. The disadvantage is the possibility of contamination with a bacterial genome (Wolbachia, etc.).



Co-funded by
the European Union

DNA barcoding

Example of Bar mtDNA sequence of a museum butterfly specimen

```
>F.696. 709 nucleotides.  
AACATTATATTTTATTTTTGGAATTTGAGCAGGTATAGTAGGAACTTCATTAAGA  
T TACTAATTCGAGCGGAATTAGGAACCCCAGGATCTTTAATTGGAGATGATCAAATTTATAATACAATTGTAACAGCCCA  
TGCATTTATTATAATTTTTTTCATAGTAATACCTATTATAATTGGAGGATTTGGAAATTGATTAGTCCCTTTAATGTTAG  
GAGCCCCTGATATAGCATTCCCACGAATAAATAATATAAGATTTTGACTTTTACCACCATCTTTAACTCTTTTAATTTCC  
AGAAGTATCGTAGAAAATGGAGCAGGAACAGGATGAACAGTTTATCCCCCCTATCTTCTAACATTGCCCATAGAGGTAG  
TTCAGTAGATCTTGCAATTTTTTCTTTACATTTAGCTGGAATTTTCATCTATTATAGGAGCAGTTAATTTTATTACCACAA  
TTATTAATATACGAATTAATAATATATCATTTCGATCAAATACCATTATTTATTTGAGCTGTGGGAATTACAGCATTTTTTA  
T TACTTTTATCTTTGCCAGTATTAGCTGGAGCTATTACTATACTTTTAACCGATCGTAATCTTAATACATCATTTTTTTGA  
TCCAGCTGGAGGAGGTGATCCTATTTTATACCAACATTTATT
```



Co-funded by
the European Union

Species identification (Blast)

NIH National Library of Medicine
National Center for Biotechnology Information

BLAST® » blastn suite » results for RID-WVMKRPPP01N

Home Recent Results Saved Strategies Help

[Edit Search](#) [Save Search](#) [Search Summary](#) [How to read this report?](#) [BLAST Help Videos](#) [Back to Traditional Results Page](#)

Job Title F.696. 709 nucleotides.

RID [WVMKRPPP01N](#) Search expires on 02-16 15:59 pm [Download All](#)

Program BLASTN [Citation](#)

Database nt [See details](#)

Query ID lcl|Query_194257

Description F.696. 709 nucleotides.

Molecule type dna

Query Length 658

Other reports [Distance tree of results](#) [MSA viewer](#)

Filter Results

Organism only top 20 will appear ☐ exclude

Type common name, binomial, taxid or group name

[Add organism](#)

Percent Identity to **E value** to **Query Coverage** to

[Filter](#) [Reset](#)

Descriptions [Graphic Summary](#) [Alignments](#) [Taxonomy](#)

Sequences producing significant alignments [Download](#) [Select columns](#) [Show](#) 100 [?](#)

☒ select all 100 sequences selected [GenBank](#) [Graphics](#) [Distance tree of results](#) [MSA Viewer](#)

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	Temnora subapicalis voucher PD-BC 760 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	Temnora subapi...	1155	1155	100%	0.0	98.33%	658	JN678591.1
<input checked="" type="checkbox"/>	Temnora griseata voucher Lope11-0902 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	Temnora griseata	1066	1066	100%	0.0	95.90%	658	MK187801.1
<input checked="" type="checkbox"/>	Temnora griseata voucher TDGABb-089 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	Temnora griseata	1061	1061	100%	0.0	95.74%	658	HQ573964.1
<input checked="" type="checkbox"/>	Temnora peckoveri cytochrome oxidase subunit 1 gene, partial cds; mitochondrial	Temnora peckov...	1044	1044	100%	0.0	95.29%	716	FJ485757.1
<input checked="" type="checkbox"/>	Temnora peckoveri voucher BC-Basq0521 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	Temnora peckov...	1044	1044	100%	0.0	95.29%	658	JN678576.1
<input checked="" type="checkbox"/>	Temnora elegans voucher MA05-06-02-49 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	Temnora elegans	1027	1027	100%	0.0	94.83%	658	JN678555.1
<input checked="" type="checkbox"/>	Temnora livida voucher TDGABb-031 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	Temnora livida	1022	1022	100%	0.0	94.68%	658	HQ573906.1
<input checked="" type="checkbox"/>	Temnora sardanus voucher BC-MNHNP0135 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	Temnora sardanus	1022	1022	100%	0.0	94.68%	658	JN678585.1
<input checked="" type="checkbox"/>	Temnora palpalis cytochrome oxidase subunit 1 gene, partial cds; mitochondrial	Temnora palpalis	1016	1016	100%	0.0	94.53%	716	FJ485756.1

nucleotide Blast:

<https://blast.ncbi.nlm.nih.gov/Blast.cgi>

highest homology:

= *Temnora subapicalis*



Co-funded by
the European Union



Siedlce University
of Natural Sciences
and Humanities

Mendel
University
in Brno



Czech University
of Life Sciences Prague

Partners:



Thank you for your attention!

This presentation has been supported by the Erasmus+ KA2 Cooperation Partnerships grant no. 2021-1-SK01-KA220-HED-000032068 "Innovation of the structure and content of study programs in the field of animal genetic and food resources management with the use of digitalisation - Inovácia obsahu a štruktúry študijných programov v oblasti manažmentu živočíšnych genetických a potravinových zdrojov s využitím digitalizácie". The European Commission support for the production of this presentation does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.



Aleš Knoll



knoll@mendelu.cz



Co-funded by
the European Union