Molecular Applications in Zoology

DNA isolation, PCR and DNA sequencing, basics in sequence data handling

Zuzana Starostová
1. Introduction, DNA isolation, PCR and DNA sequencing (14.2.2023, lecturer Zuzana Starostová)
2. Next generation sequencing – introduction, short-read and long-read sequencing methods, assemblies (21.2.2023, lecturer Radka Reifová)
3. How to read and make phylogenetic trees (28.2.2023, lecturer Zuzana Starostová)
4. Use of molecular phylogenetics in zoology (7.3.2023, lecturer Zuzana Starostová)
5. Microsatellites and molecular identification of species, individuals and sex (14.3.2023, lecturer Pavel Munclinger)
6. Population structure and paternity studies (21.3.2023, lecturer Pavel Munclinger)
7. Molecular phylogeography (28.3.2023, lecturer Pavel Hulva)
8. Next generation sequencing – applications: whole genome sequencing, exome sequencing, ddRAD sequencing, metagenomics (4.4.2023, lecturer Radka Reifová)
9. Gene expression – quantitative PCR, expression microarrays, transcriptome sequencing and analysis, epigenomics (11.4.2023, lecturer Radka Reifová)
10. Functional genetic variability: From SNP to selection (18.4.2023, lecturer Michal Vinkler)
11. Cytogenetic methods and their application in zoology (25.4.2023, lecturer František Šťáhlavský)

Practicals:
Day 1: database search and basic DNA sequence data handling

Day 2: lab work – molecular method of sex determination in birds – DNA isolation, PCR, gel electrophoresis
Exam:
written test covering content of the lectures

Literature:
Molecular techniques and zoology

- since 90s molecular techniques more available - automatization, lower costs
- commercial synthesis of oligonucleotides, sequencing services, user friendly kits
- routine application of molecular techniques into fields like ecology, systematics, ethology, conservation biology...
- new methods – new resources - new questions
Nucleic acids

- Nucleic acids are macromolecules that store genetic information and enable protein production
- DNA and RNA – molecules composed of strands of nucleotides
- DNA is composed of a phosphate-deoxyribose sugar backbone and the nitrogenous bases adenine (A), guanine (G), cytosine (C), and thymine (T).
- RNA has ribose sugar and the nitrogenous bases A, G, C, and uracil (U).

More about RNA in lesson # 9

Let’s talk about DNA
DNA

- is carrying the genetic instructions used in the growth, development, functioning, and reproduction of all known living organisms

**Structure:**
- two polynucleotide strands
- the nitrogenous bases of the two separate polynucleotide strands are bound together, according to base pairing rules (A with T and C with G), with hydrogen bonds to make double-stranded DNA
- the nitrogenous bases - two types: 
  - **pyrimidines** are thymine and cytosine; 
  - **purines** are adenine and guanine
- two strands of DNA run in opposite directions to each other and are thus **antiparallel**
Sources of DNA for sampling:

- Destructive sampling – tissue (fresh, frozen, dry, preserved in ethanol) – ideal spleen, muscle- e.g. tongue

- Invasive but non-destructive sampling- buccal smears, fin clips, tail tips (rodents), blood sample

- Non-invasive sampling (DNA collected without capturing or usually even touching an animal)- feces, hair, feathers, saliva, shed skin, eggshell
  - DNA yield and quality tend to be lower (higher risk of contamination)
Figure 1. Non-disruptive DNA sampling methods, and their overlaps with non-invasive and non-lethal sampling methods.

fur traps

- different for large vs. small species
- effectiveness increased by attractants - cats - *Valeriana officinalis*, deers - anise
Fig. 1 Materials and application of the optimized hair trap. The trap consists of a PVC pipe with two holes, an iron nail, and a plastic roll (a). The plastic roll is wrapped with double-faced adhesive tape and attached inside the pipe by the nail (b). Traps are placed in entrances of burrows (e), and as an animal passes, hairs are plucked by the rolling tape (d). Most of plucked hair contained bulbs with a sufficient amount of follicular tissue (e and f).

An optimized hair trap for non-invasive genetic studies of small cryptic mammals

Tobias Erik Reiners · Jorge A. Encarnação · Volkmann Wolters
eDNA

Environmental DNA or eDNA is DNA that is collected from a variety of environmental samples (soil, water, snow, air) rather than directly sampled from an individual organism.

cave-dwelling amphibian *Proteus anguinus* inhabits subterranean waters of the north-western Balkan Peninsula. Because only fragments of its habitat are accessible to humans, this endangered salamander’s exact distribution has been difficult to establish.
Loch Ness Monster may be a giant eel, say scientists

5 September 2019

Of the common theories associated with the 1,000 or so sightings of something swimming in the water at Loch Ness, the environmental DNA data obtained suggests at least one theory remains plausible.

Eels returned the largest proportion of DNA from the 250 water samples taken throughout Loch Ness. Typically not gigantic, could an extremely large European eel be the creature people have seen moving "like a torpedo" in the water? The data obtained suggests this may be possible, although no eel of the size described in some accounts has ever been caught or found.

Infrequent visitors such as seals and possibly sturgeons may account for some sightings, but wales, standing waves and logs are the basis of most.
DNA Extraction from Dry Museum Beetles without Conferring External Morphological Damage

M. Thomas P. Gilbert\textsuperscript{1,2*}, Wendy Moore\textsuperscript{3}, Linea Melchior\textsuperscript{1,4}, Michael Worobey\textsuperscript{1}

The Isolation of Nucleic Acids from Fixed, Paraffin-Embedded Tissues—Which Methods Are Useful When?

M. Thomas P. Gilbert\textsuperscript{1*}, Tamara Haselkom\textsuperscript{1}, Michael Bunce\textsuperscript{2}, Juan J. Sanchez\textsuperscript{3}, Sebastian B. Lucas\textsuperscript{4}, Laurence D. Jewell\textsuperscript{5}, Eric Van Marck\textsuperscript{6}, Michael Worobey\textsuperscript{1}
Fossils as a source of DNA – ancient DNA

- authentic ancient DNA - up to several hundred thousand years old samples
- bones, artificially or naturally mummified animal remains, remains in permafrost, fossil eggs, paleofaeces
- extremely high risk of contamination- specialized laboratories
- typically short fragments of DNA, but now also whole genomes (e.g.: mt genomes of mammoth, moas, whole genome of the Neanderthal)

DNA of extinct birds extracted from ancient eggshell

By Pallab Ghosh
Science correspondent, BBC News

Researchers have found that eggshells of extinct bird species are a rich source of preserved DNA.

An international team isolated the delicate DNA molecules of species including the massive "elephant birds" of the genus Aepyornis.

The Proceedings of the Royal Society B research demonstrated the approach also on emu, ducks and the extinct moa.

The team says that the technique will enable researchers to learn more about ancient birds and why they died out.
Storing samples for DNA isolation:

- frozen (–20°C usually OK, -80°C for long periods)
- dry
- fixed tissue in 96% ethanol (formaldehyde not ideal, but not impossible)
- samples for RNA isolation can be stored in RNAlater (or -80°C)
Genetic banks:

Biorepository: Department of Zoology, Charles University in Prague

Institution Information
DNA isolation

- different methods and technologies are available
- all methods involve disruption and lysis of the starting material followed by the removal of proteins and other contaminants and finally recovery of the DNA
- removal of proteins - proteinase K
- choice of a method depends on: the required quantity of the DNA, the purity required for downstream applications, and the time and costs

Examples of methods:
- organic extraction – phenol/chloroform
- specialized kits usually using binding of the DNA to a solid-phase support
Different genes or genomes = different questions

nuclear DNA
-biparental inheritance in diploid sexually reproducing organisms
-nuclear coding genes - lower mutation rate (in mammals mtDNA cca 10x higher mutation rate)
-microsatellite loci - mutate much more rapidly than most other types of sequences
-sex chromosomes - in mammals, the Y chromosome is the paternally inherited - most of it does not undergo recombination - useful marker for retracing male lineages

Location of mtDNA and nuclear DNA in an animal cell. The mitochondria of an animal cell harbor maternally transmitted mtDNA. Nuclear DNA, which is transmitted by both parents, is found in the nucleus, and contains repeated elements termed microsatellites.
© 2012 Nature Education
Different genes or genomes = different questions

**mitochondrial DNA**
- uniparentally inherited - maternal (but exceptions exist)
- mtgenomes relatively small - between 15,000 and 17,000 bp
- structure, size and arrangement of genes are relatively conserved
- no recombination = offspring will usually have exactly the same mitochondrial genome as its mother
- the mutation rate is relatively high
- non-coding control region (incl. D-loop), evolves rapidly in many taxa
- small effective population size – sensitive to demographic events such as bottlenecks
- available universal primers

![Figure 2.1](image_url)  
Typical gene organization of vertebrate mtDNA. Unlabelled dark bands represent 22 transfer RNAs (tRNAs). Gene abbreviations starting with ND are subunits of NADH dehydrogenase and those starting with CO are subunits of cytochrome c.
Polymerase Chain Reaction = PCR

• method for rapid amplification of DNA sequences – selected fragment of DNA

• amplified fragment of DNA is defined by pair of primers

• Primer - short DNA fragments containing sequences complementary to the target region

```
primer       AGGGGACGTACACTCAGCTTT
template DNA TCCCCTGCATGTGAGTCGAAA
```

fragment selected for amplification
Polymerase Chain Reaction = PCR

- developed in mid 80s by Kery Mullis
- 1993 – Kery Mullis and Michael Smith awarded the Nobel Prize in Chemistry for their work on PCR

PCR uses repeated cycles of heating and cooling to make many copies of a specific region of DNA

3 steps
1. temperature is raised to near boiling - causing the double-stranded DNA to separate (denature) into single strands
2. temperature is decreased and short DNA sequences (primers) bind (anneal) to complementary matches on the target DNA sequence
3. at a slightly higher temperature the enzyme Taq polymerase binds to the primed sequences and adds nucleotides to extend the second strand
**Initial Denaturation of DNA**

- **94 - 98°C**
- **3 min 30 s**
- **Denaturation**
  - Temperature is increased to separate DNA strands

**Denaturation**

- **50 - 65°C**
- **30 - 45 s**
- **Annealing**
  - Temperature is decreased to allow primers to base pair to complementary DNA template

**Extension**

- **72°C**
- **45 s**
- **Prolonged Extension in the Last Cycle**
- **3 min 30 s**
- **Extension**
  - Polymerase extends primer to form nascent DNA strand

*Extension time depends on the length of amplified fragment - nucleotides are incorporated in the rate cca 35 nucleotides per second*

**Exponential Amplification**

- **1st cycle** → **2nd cycle** → **3rd cycle** → **4th cycle** → **30th cycle**
- **2^1 = 2 billion copies**
- **Process is repeated, and the region of interest is amplified exponentially**

*Source: New England Biolabs*
Polymerase Chain Reaction = PCR

PCR requires:
- two specific oligonucleotide primers
- thermostable DNA polymerase - Taq polymerase
- dNTP’s
- template DNA
- MgCl₂
- buffer (supplied with the polymerase)
- water

PCR is performed in the thermal cycler (also known as a thermocycler or PCR machine)

PCR usually prepared in small volumes – 20-100 µl
PCR primers

- short oligonucleotides 17-30bp long
- sequences complementary to the target region
- no repetitive sequences
- no sequence complementarity within or between primers – secondary structures, primer dimer
- annealing temperature dependent upon primer sequence (~ 50% GC content)
- software for primer design (web tools)
RT-PCR (reverse transcription PCR) and Real-Time) PCR

RT-PCR – method to detect RNA expression
- modification of PCR, RNA is template
- qualitatively detect gene expression through creation of complementary DNA (cDNA) transcripts from RNA

Real-Time PCR (also qPCR)
- quantitatively measure the amplification of DNA using fluorescent probes
- specially modified cycler
Gel electrophoresis

method for separation and analysis of macromolecules (DNA, RNA, proteins) based on their size and charge

electrophoresis = transport by electricity

-DNA is negatively charged – fragments of different length are separated by applying an electric field - they move through a gel (usually agarose gel, for more precise separation polyacrylamide gel)

-shorter molecules move faster and migrate farther than longer ones
**Gel electrophoresis**

- Electric field consists of a negative charge at one end which pushes the molecules through the gel, and a positive charge at the other end that pulls the DNA.

- Molecules being sorted are dispensed into a well in the gel material.

- Gel is placed in an electrophoresis chamber, which is then connected to a power source.

- When the electric current is applied, fragments molecules move.

- The different sized molecules form distinct bands on the gel.
Gel electrophoresis
- electrophoresis chamber is filled with buffers
  - provide ions that carry a current and maintain stable pH

- most common buffers for nucleic acids
  Tris/Acetate/EDTA (TAE) or Tris/Borate/EDTA (TBE) buffers

- samples are mixed with dye for monitoring

- to visualize DNA fragments - ethidium bromide bind to the DNA - visible under UV (SYBR Green also used)

- size standard (ladder) for estimating size of fragments
ELFO of total genomic DNA
ELFO after PCR

Size standard (ladder)
RNA gel electrophoresis
DNA sequencing:

Sanger Method for DNA Sequencing = dideoxynucleotide sequencing
method uses 2',3'-dideoxynucleotide triphosphates (ddNTPs), molecules that differ from deoxynucleotides by having a hydrogen atom attached to the 3' carbon rather than an OH group

-ddNTPs terminate DNA chain elongation because they cannot form a phosphodiester bond with the next deoxynucleotide

Sequencing reaction:
modified PCR reaction: purified PCR product (template), primer (one), polymerase, fluorescently labelled dideoxyNTPs, normal deoxynucleosidetriphosphates (dNTPs)
In the sequencer:

1. **capillary electrophoresis** – size separation of fragments
2. **laser detection** - each of the four dideoxynucleotide chain terminators is labeled with fluorescent dyes, each of which emit light at different wavelengths

http://www.wellcome.ac.uk/Education-resources/Teaching-and-education/Animations/DNA/WTDV026689.htm
1. Reaction mixture
   - Primer and DNA template
   - DNA polymerase
   - ddNTPs with fluorochromes
   - dNTPs (dATP, dCTP, dGTP, and dTTP)

2. Primer elongation and chain termination

3. Capillary gel electrophoresis separation of DNA fragments

4. Laser detection of fluorochromes and computational sequence analysis

from Wikipedia
read length and quality:
currently read lengths of approximately up to 800–1000 bp, longer fragments need to be sequenced from both sides (both strands of DNA):

problems of DNA sequencing with the Sanger method:

• poor quality in the first 15-40 bases of the sequence due to primer binding

• deteriorating quality of sequencing traces after 700-900 bases
Sanger DNA sequencing:

• automated sequencing machines (sequencers) – different number of capillaries: 4, 8, 16, 24, ...

• specializes sequencing laboratories (services)
Chromatogram files:
Can be opened in different software - Chromas, BioEdit, DNASTAR, ...
other sequencing methods – next generation sequencing

**illumina**

**Ion Torrent**

... more on this topic in special lectures