

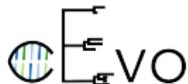
Computational Biology

Lecturers:
Tanja Stalder, Carsten Magnus & Tim Vaughan

Teaching Assistants:
Jūlija Pečerska, Jérémie Sciré,
Sarah Nadeau & Marc Manceau

Computational Evolution
Department of Biosystems Science and Engineering

HS 2019



- 01a-Introduction
 - Teaching staff
 - Logistics
 - Computational Biology
 - Brief History of evolution
 - Course overview
- 01b-Sequencing techniques
 - About sequences
 - What are sequences?
 - How can we obtain the sequences?
 - Sequencing methods
- References

Teaching staff

Lecturers



Tanja
Stadler



Carsten
Magnus



Tim
Vaughan

Tutors



Jūlija
Pečerska



Jérémie
Sciré



Sarah
Nadeau



Marc
Manceau

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Lectures and tutorials

► 13 lectures

- every Monday from 23 September - 16 December in Zurich and Basel
- 3.15pm – 5pm (15min break in between)
- lectured in one location, live-broadcast in teleconference style to the other location

► 13 tutorials - pen+paper exercises, discussion of homework

- Zurich: every Monday 5.15pm – 6pm
- Basel: every Thursday 12.00pm – 12.45pm

► Locations:

- Please refer to
<https://bsse.ethz.ch/cevo/education/cb-materials.html>

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Exam – 75% of grade

► Written exam:

- counts for 75% of final grade
- length 90min
- during examination period in Jan./Feb. 2020, date will be announced by ETH
- both theoretical concepts and biology will be tested (pen and paper exam)
- exam preparation: questions at end of each lecture, short briefing at the end of the lecture series, and question session in Jan. 2020

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Homework – 25% of final grade

- ▶ There will be 5 graded homework assignments:
 - counts for 25% of final grade
 - lecture-related algorithms and accompanying theoretical questions
 - coded in R
 - ☞ special lesson on how to complete homework assignments in first tutorial
 - handed out every other week (30 Sep, 14 Oct, 28 Oct, 11 Nov, 25 Nov), accompanied by a related pen and paper exercise in the tutorial of that week
 - on the week following the assignments, the TAs will discuss the previous homework and will provide help with the current homework
- ▶ Assignment rules:
 - working in groups is okay
 - however, you must submit your **own** code
 - code will be tested for plagiarism (e.g. simple renaming of variables will be detected)
 - assignment published on Mondays via Moodle, solutions have to be submitted by 12:00h CET on Monday two weeks later via Moodle

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Moodle

- ▶ All CB Material will be made available on Moodle: <https://moodle-app2.let.ethz.ch/course/view.php?id=11222>
 - slides
 - video recordings
 - lecture notes (β -version; please report any typos, errors, or suggestions for improvement! The student(s) with the most valuable feedback will be rewarded at the end of the course.)
 - additional references
 - homework assignments: download and upload
 - discussion forums

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Questions and feedback

► Feedback:

- Please tell/ask/write to us whenever something needs improvement via
 - ☞ the feedback forum on Moodle, or
 - ☞ the anonymous feedback tool <https://www.bsse.ethz.ch/cevo/education/cb-materials.html>.

► Questions:

- Moodle forums: please help each other and answer questions!
- Moodle forums: additionally curated by the tutors, first place to go for help from tutors, lecturers.
- For personally sensitive matters, write to us directly to set up a meeting.
- In class: please ask and interact!

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Computational Biology.

01a-Introduction

- Teaching staff
- Logistics
- Computational Biology
- Brief History of evolution
- Course overview

01b-Sequencing techniques

- About sequences
- What are sequences?
- How can we obtain the sequences?
- Sequencing methods
- References

What is Computational Biology?

“Bioinformatics is a young, rapidly evolving, and interdisciplinary research field. It develops and applies computational techniques and processes to analyze genetic and other biological information and to tackle challenging problems in biology.”

<http://www.cbb.ethz.ch/>

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Focus of this course

A revolution in sequencing technologies makes genetic sequences one of the richest data sources for answering biological questions. In this course, we will discuss how to obtain and analyze genetic sequences:

- ▶ **Obtaining & Organizing Sequences**

How do we obtain sequences, align them, and what does data mining tell us about them?

- ▶ **Molecular Evolution**

How does the genetic information of individuals change through time?

- ▶ **Phylogenetics**

How are the individuals related?

- ▶ **Phyldynamics**

What are the population dynamics giving rise to the individuals in the phylogeny?

[01a-Introduction](#)

[Teaching staff](#)

[Logistics](#)

[Computational Biology](#)

[Brief History of evolution](#)

[Course overview](#)

[01b-Sequencing techniques](#)

[About sequences](#)

[What are sequences?](#)

[How can we obtain the sequences?](#)

[Sequencing methods](#)

[References](#)

Sequences tell us about macroevolution

Macroevolution: individuals = species

► **Obtaining & Organizing Sequences**

Alignment of *homolog* sequences from different species.

► **(Molecular) Evolution**

(Genetic information of) species changed through time.

► **Phylogenetics**

Phylogeny displays species relationships.

► **Phylogenetics**

Population dynamics is the speciation and extinction process.

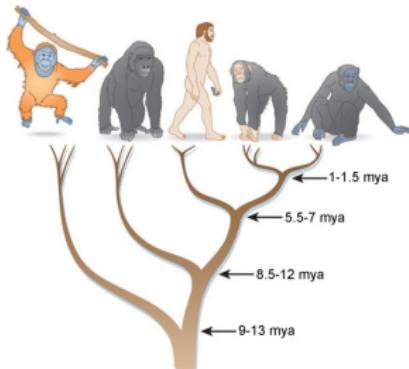


Figure adapted from [Paabo, 2003]

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Further key applications

► **Epidemiology** - individuals = infected hosts

- Phylogeny displays transmission history.
- Population dynamics are the transmission and becoming non-infectious processes.

► **Immunology** - individuals = B cells

- Phylogeny displays B cell differentiation (through somatic hypermutation).
- Population dynamics is the B cell generation and loss process.

► **Cancer** - individuals = cells

- Phylogeny displays relationship of different cancer cells and healthy cells.
- Population dynamics is the spread and loss of cell types.

► **Languages** - individuals = languages

- Phylogeny displays language evolution.
- Population dynamics is the gain and loss of languages.

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Sequences differ due to evolution

"Evolution is change in the heritable traits of biological populations over successive generations. Evolutionary processes give rise to diversity at every level of biological organisation, including the levels of species, individual organisms, and molecules." (Wikipedia)

How does evolution occur?

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Brief History of evolution.

01a-Introduction

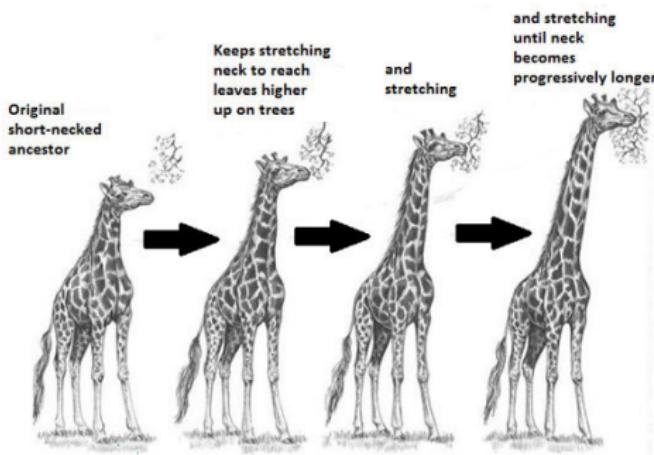
- Teaching staff
- Logistics
- Computational Biology
- Brief History of evolution**
- Course overview

01b-Sequencing techniques

- About sequences
- What are sequences?
- How can we obtain the sequences?
- Sequencing methods
- References

Lamarckian evolution

Evolution through use and disuse of features (Lamarck, 1809):



Lamarck's famous Giraffe example

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

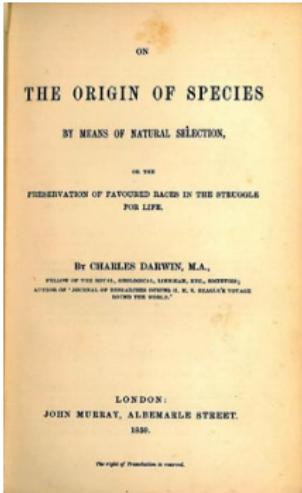
Sequencing methods

References

Darwin's theory of evolution

Evolution by the means of natural selection occurs given:

- ▶ **Multiplication** of individuals
- ▶ **Variation** in the phenotype of individuals
- ▶ **Heredity** of the phenotype
- ▶ **Competition** between phenotypes to survive and multiply



What encodes the phenotype and is inherited?

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Genetics

Gregor Mendell, an Austrian monk, crossed thousands of pea plants with white and purple flowers. His observations revealed the mechanisms of heredity (1865):

- ▶ **gene:** entity encoding a phenotype (e.g. flower colour of peas).
- ▶ **allele:** version of a gene (one allele of the colour gene may encode for white (y); another allele for purple (Y)).
- ▶ **genotype:** the collection of genes of one individual.

- ▶ Each individual has two alleles for each gene, a random one from the father and a random one from the mother.
- ▶ Dominant allele is the one determining the phenotype (e.g. Yy peas are purple, i.e. Y is dominant); the other allele is the recessive allele.

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Modern Synthesis (also called Neo-Darwinism)

Term coined by [Huxley et al., 1942]:

- ▶ Combine Darwin's natural selection with Mendelian genetics
(population genetics: Fisher, Wright, Haldane, around 1930)

Modern synthesis became key working hypothesis as a mechanism for evolution:

- ▶ Modern Synthesis explains natural population dynamics
(Theodosius Dobzhansky)
- ▶ Bridge from Modern Synthesis to Paleontology
(George Simpson)

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

DNA encodes the genes

- ▶ DNA (Deoxyribonucleic acid) first isolated by Friedrich Miescher (born in Basel in 1844; Institute here in Basel named after him).
- ▶ DNA is a double helix: published in 1953 by Watson & Crick; based on images by Rosalind Franklin; Nobel prize in 1962.



This figure is purely diagrammatic. The two ribbons symbolize the two phosphate-sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis.

Figure adapted from
[Watson and Crick, 1953]

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

DNA helix

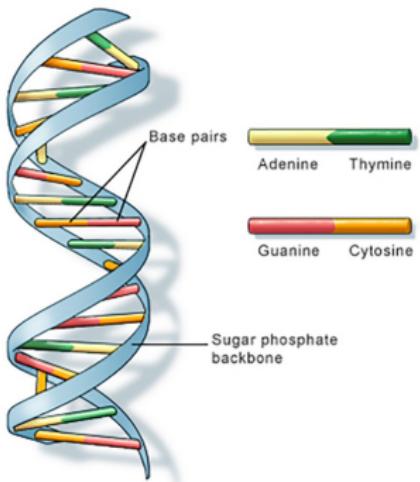


Figure adapted from [ghr.nlm.nih.gov, 2015]

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

DNA is characterized by a string of the letters A, T, G, C. We call this string the **genetic sequence**.

Central dogma in biology

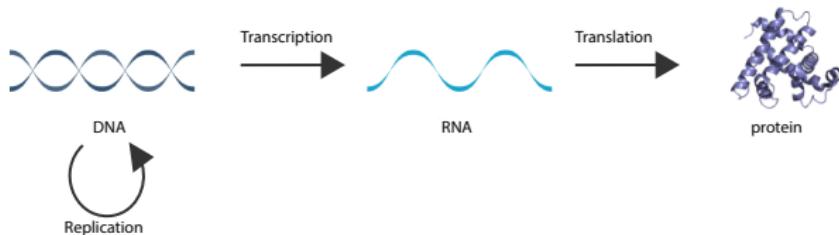


Figure adapted from [adtbio.com, 2015]

The **genotype** determines the **phenotype**!

Richard Dawkins:

- genotype is recipe, phenotype is cake

John Maynard Smith:

- genotype is plan for building an airplane, phenotype is airplane

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

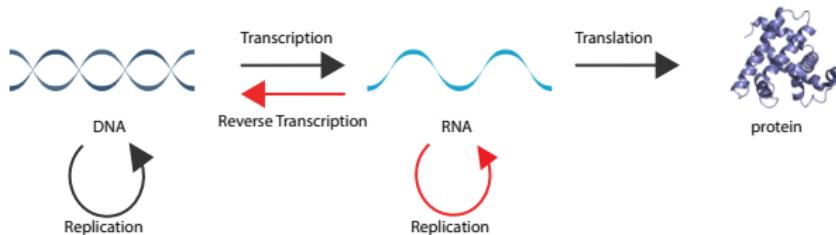
What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Central dogma in biology - no rule without exceptions!



01a-Introduction

- Teaching staff
- Logistics
- Computational Biology
- Brief History of evolution
- Course overview

01b-Sequencing techniques

- About sequences
- What are sequences?
- How can we obtain the sequences?
- Sequencing methods

References

Discovery of reverse transcription by Howard Temin and David Baltimore was awarded with Nobel prize in Physiology or Medicine, 1975.

DNA replication

During DNA replication, DNA may change due to:

- ▶ point mutations
- ▶ recombination
- ▶ insertions
- ▶ deletions

This variation in DNA causes the variation in phenotypes.

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Darwin today

- ▶ **Multiplication** (replication) of DNA leads to offspring. DNA (genotype) determines the phenotype.
- ▶ **Variation** in the offspring phenotypes occurs due to mutations, recombination, insertions, deletions in DNA.
- ▶ **Heredity** of the phenotype occurs due to the DNA being passed on, and the phenotype being encoded by the genotype.
- ▶ **Competition** between phenotypes to survive and multiply occurs.

This view neglects any impact of the environment. Epigenetics is suggested to be a mechanism to transmit environmentally acquired phenotypes (loops back to Lamarck)!

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

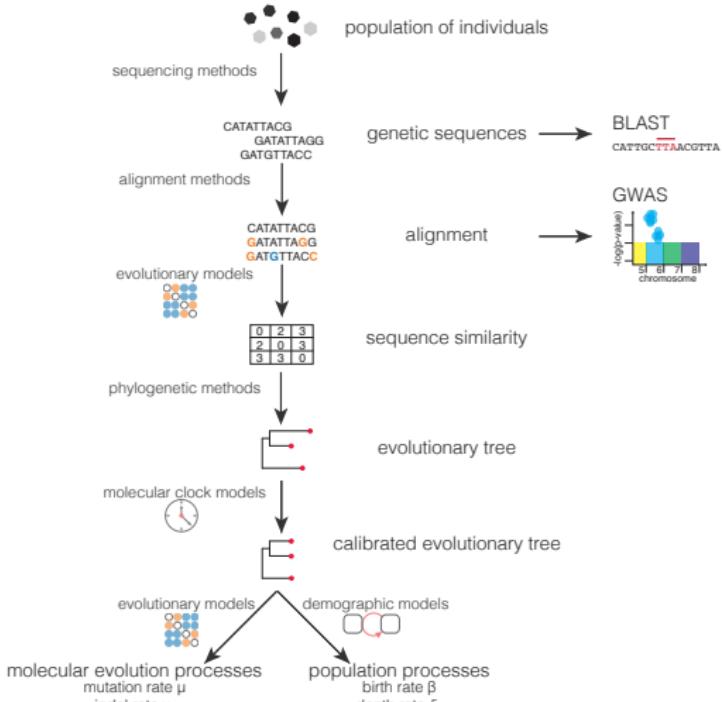
What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Why, how, and what can we learn from genetic sequences?



01a-Introduction

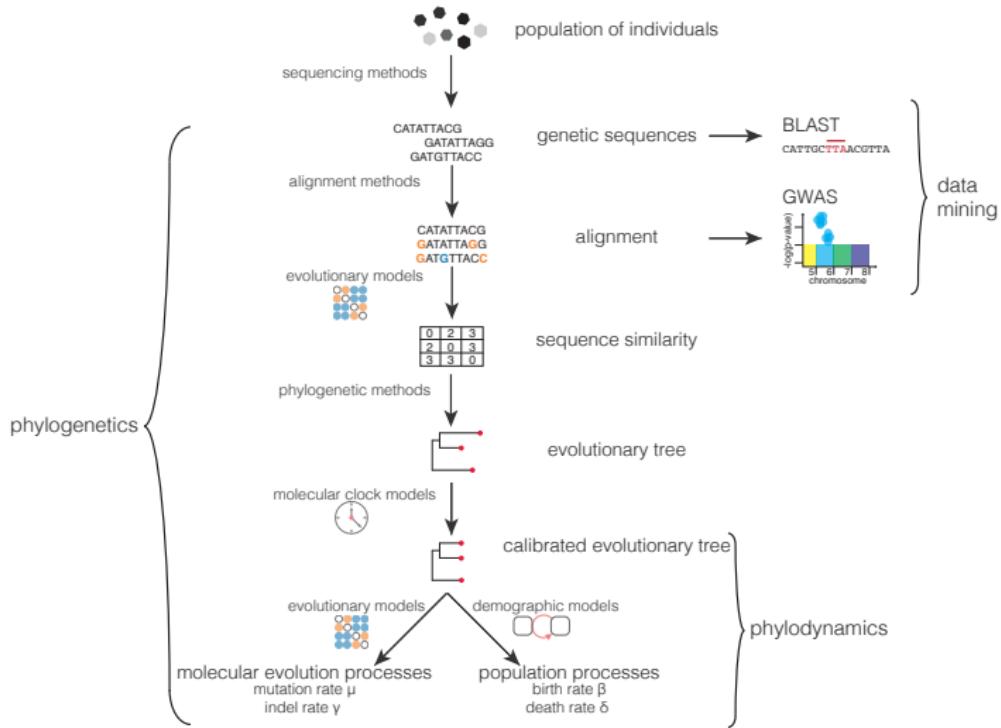
Teaching staff
Logistics
Computational Biology
Brief History of evolution
Course overview

01b-Sequencing techniques

About sequences
What are sequences?
How can we obtain the sequences?
Sequencing methods

References

Why, how, and what can we learn from genetic sequences?



01a-Introduction

- Teaching staff
- Logistics
- Computational Biology
- Brief History of evolution
- Course overview

01b-Sequencing techniques

- About sequences
- What are sequences?
- How can we obtain the sequences?
- Sequencing methods

References

Learning goals

In this course you will learn how to

- ▶ obtain sequences
- ▶ use sequences for identification purposes (BLAST) and obtaining correlations between genetic variants and certain phenotypes (GWAS); (data mining)
- ▶ construct phylogenetic trees based on sequences and interpret them (phylogenetics)
- ▶ use phylogenetic trees in order to study population dynamic process (phylodynamics)
- ▶ design a genetic sequence analysis
- ▶ derive models with pen and paper
- ▶ implement computational algorithms to solve problems involving repeated tasks or big data

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Reading the genetic code

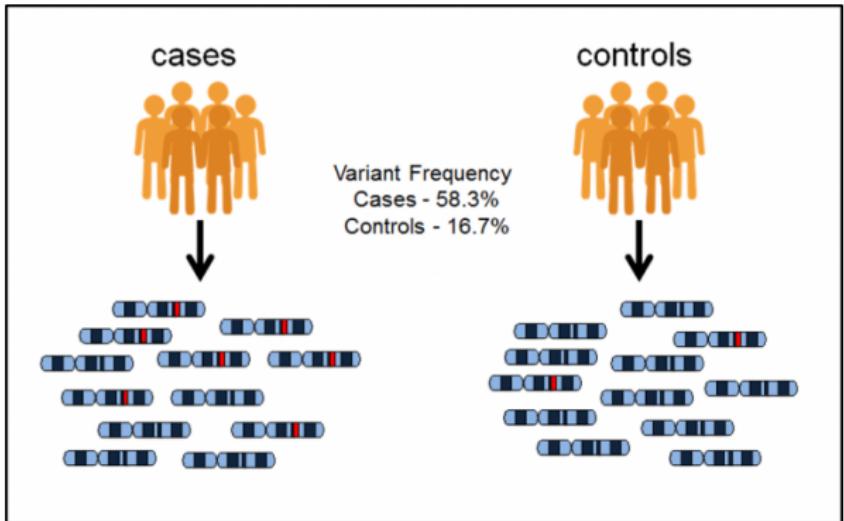
- ▶ Late 1990s: race between universities (led by Francis Collins) and Celera Genomics (led by Craig Venter) to sequence the human genome.
- ▶ June 2000: Collins & Venter announced the sequencing success. "Today we are learning the language in which God created life," President Clinton said.
- ▶ February 2001: Venter et al. publish the draft genome in Science; Collins et al. publish the draft genome in Nature.



01a-Introduction
Teaching staff
Logistics
Computational Biology
Brief History of evolution
Course overview

01b-Sequencing techniques
About sequences
What are sequences?
How can we obtain the sequences?
Sequencing methods

References

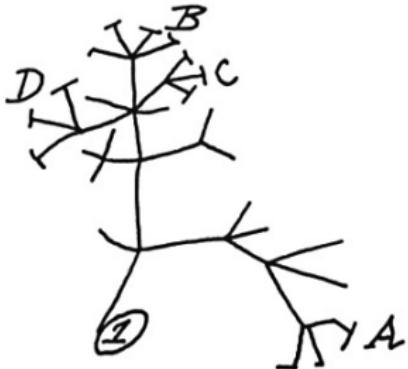


[www.ebi.ac.uk,]

Basis of an Genome Wide Association study.

- ▶ For each position in the genetic code, the distribution of characters found in the “case” group is compared with the distribution in the “control” group.

Phylogenetics



The first known sketch of an evolutionary tree
(Darwin, from his 1837 Notebook)

- ▶ Sequences are not independent samples but share an evolutionary history, the phylogenetic tree.
- ▶ Similar DNA inciates close relatives; computational tool by [Fitch and Margoliash, 1967] in Science.
- ▶ Felsenstein revolutionized statistical phylogenetic tree inference in 1980s (for an overview see [Felsenstein, 2004]).

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

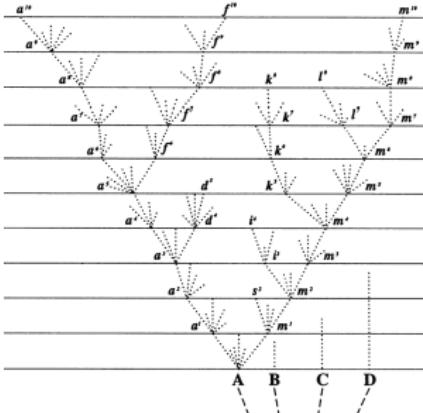
What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Phylogenetics



The first known sketch of a time tree
 (the only figure in the *Origin of Species* by Darwin, 1859)

- ▶ Determine the population dynamics (such as speciation and extinction) from the timing of events in the evolutionary tree.
- ▶ Started in the 1990s in Oxford for macroevolution and really took off for epidemiology with the publication of [Grenfell et al., 2004].

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

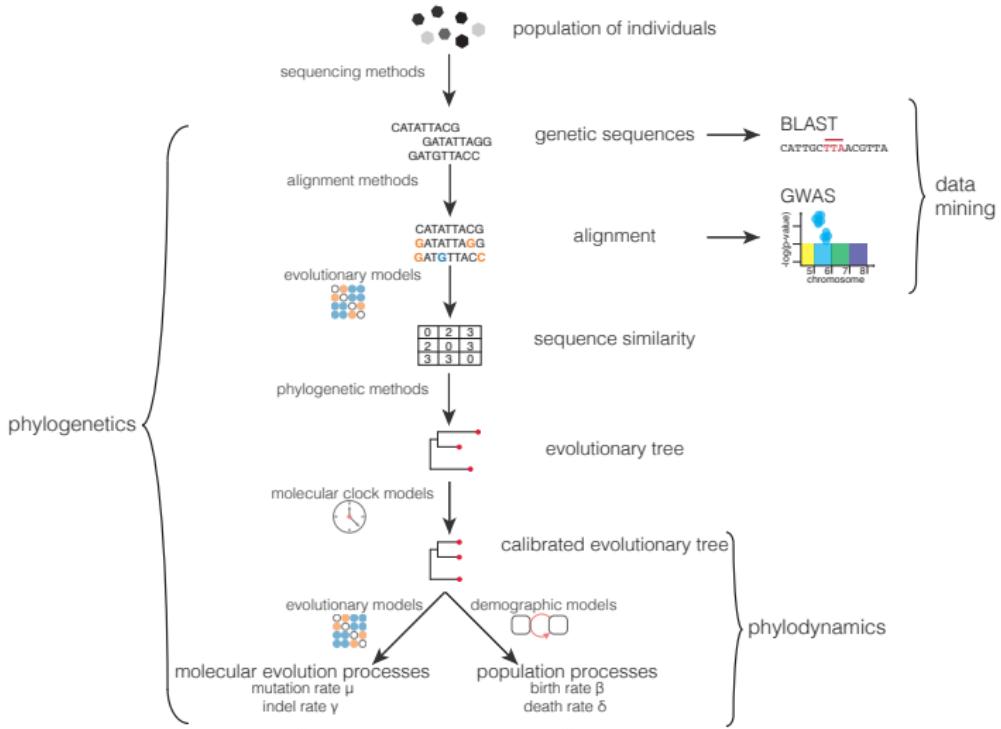
What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Why, how, and what can we learn from genetic sequences?



01a-Introduction

- Teaching staff
- Logistics
- Computational Biology
- Brief History of evolution
- Course overview

01b-Sequencing techniques

- About sequences
- What are sequences?
- How can we obtain the sequences?
- Sequencing methods

References

Sequencing.

01a-Introduction

- Teaching staff
- Logistics
- Computational Biology
- Brief History of evolution
- Course overview

01b-Sequencing techniques

- About sequences
- What are sequences?
- How can we obtain the sequences?
- Sequencing methods
- References

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

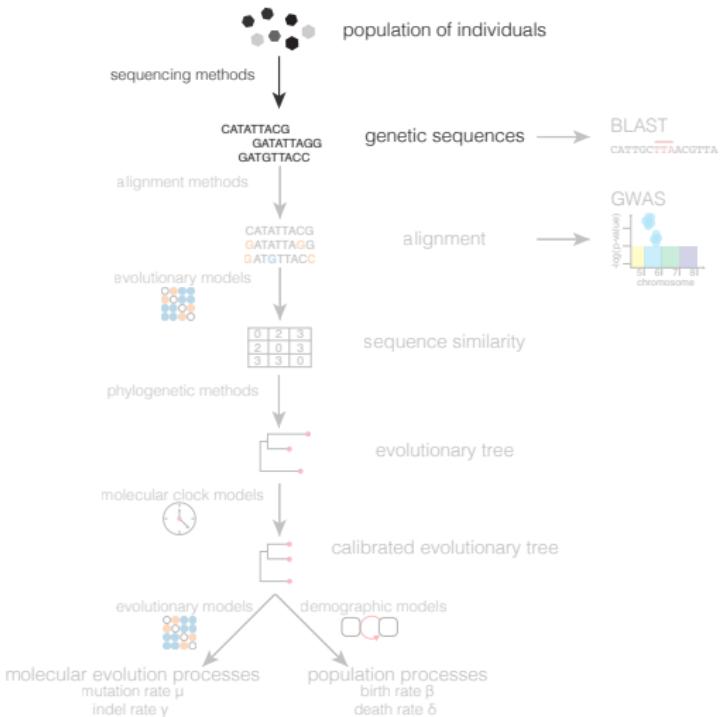
About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References



Outline of lecture 1b

- ▶ What are sequences?
- ▶ Why do we want to know the sequences?
- ▶ How can we obtain the sequences?
 - ▶ Sequencing methods
 - ▶ Sanger sequencing
 - ▶ Second-generation sequencing
 - ▶ Third-generation sequencing

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

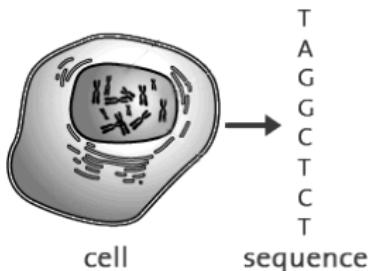
How can we obtain the sequences?

Sequencing methods

References

Sequence in our context...

Each cell of an organism contains the analogue blueprint of the genetic information of this organism.



This blueprint

- ▶ can be written in DNA or RNA language
 - ▶ contains the code for mRNA/miRNA/rRNA/.../proteins/other (useful) stuff
 - ▶ is subject to change (mutation)
 - ▶ contains the footprint of evolution of life
- ⇒ we want to read this blueprint

A sequence is an excerpt from this blueprint.

01a-Introduction
Teaching staff
Logistics
Computational Biology
Brief History of evolution
Course overview
01b-Sequencing techniques
About sequences
What are sequences?
How can we obtain the sequences?
Sequencing methods
References

Single sequence

- ▶ Read parts of the genome (DNA/RNA) of an organism
- ▶ If the surrounding sequence is known:
SNP or short indels (= insertions and deletions) of interest
(☞ lecture 2)
- ▶ Else: meaningless without context (i.e. coordinates)

To do something more meaningful:

⇒ Assemble and align sequences (☞ lecture 2)

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Step-by-step guide to obtaining a sequence



Figure adapted from [Anthro.palomar.edu, 2015]

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Step-by-step guide to obtaining a sequence



cell

Figure adapted from [Anthro.palomar.edu, 2015]

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Step-by-step guide to obtaining a sequence

1 isolate cells

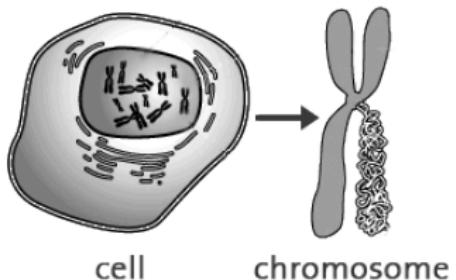


Figure adapted from [Anthro.palomar.edu, 2015]

Step-by-step guide to obtaining a sequence

- 1 isolate cells
- 2 and purify DNA/RNA from the cell (or rather cells)

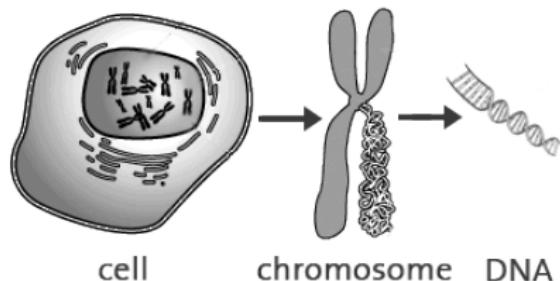


Figure adapted from [Anthro.palomar.edu, 2015]

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Step-by-step guide to obtaining a sequence

- 1 isolate cells
- 2 and purify DNA/RNA from the cell (or rather cells)
- (3) multiply (target part of) DNA/RNA using polymerase chain reaction (PCR)

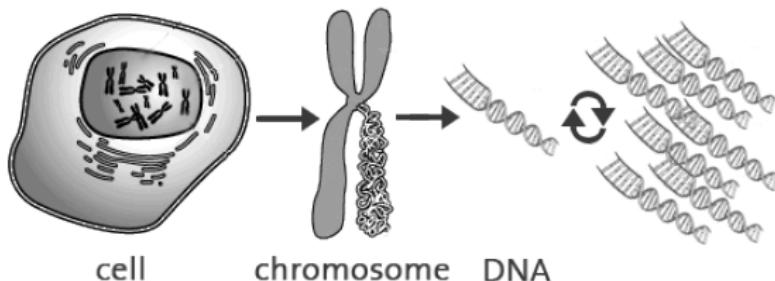


Figure adapted from [Anthro.palomar.edu, 2015]

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Step-by-step guide to obtaining a sequence

- 1 isolate cells
- 2 and purify DNA/RNA from the cell (or rather cells)
- (3) multiply (target part of) DNA/RNA using polymerase chain reaction (PCR)
- 4 apply your preferred sequencing method

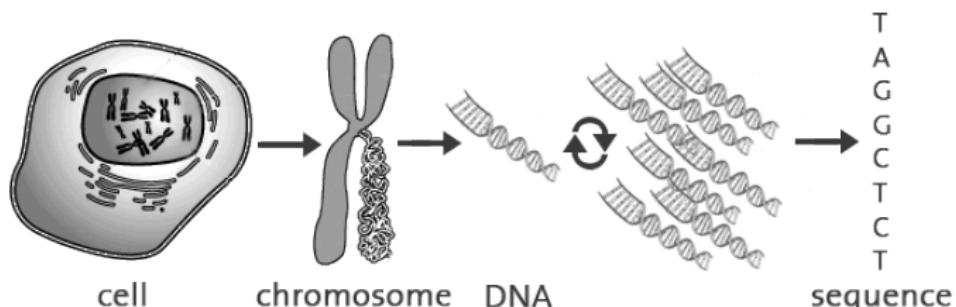


Figure adapted from [Anthro.palomar.edu, 2015]

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

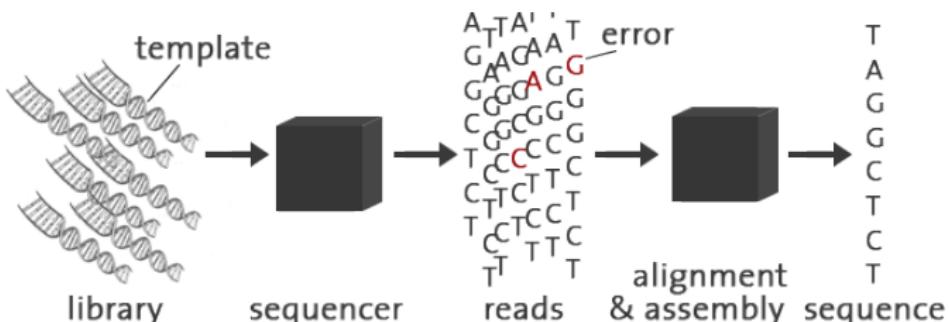
How can we obtain the sequences?

Sequencing methods

References

Step-by-step guide to obtaining a sequence

- 1 isolate cells
- 2 and purify DNA/RNA from the cell (or rather cells)
- (3) multiply (target part of) DNA/RNA using polymerase chain reaction (PCR)
- 4 apply your preferred sequencing method



01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Vocabulary

- ▶ **Template** - something that serves as a “model” for others to copy (e.g. DNA of a cell being template for amplification/sequencing)
- ▶ **Library** - mixture of different templates, ready for sequencing
- ▶ **Sequence** - order of building blocks (nucleotides) in a template, can but does not have to be a full genome
- ▶ **Sequencing** - determining the order of individual building blocks of the sequence
- ▶ **Sequencing run** - an operation round of the sequencing machine
- ▶ **Sequencing error** - difference between the sequence retrieved via sequencing and the template sequence
- ▶ **Read** - an output from a sequencing run - just a small part of the full genome
 - ▶ can contain sequencing errors (and often does)
- ▶ **Assembly** - aligning and merging reads in order to reconstruct the original sequence of the template

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Overview of sequencing methods

- ▶ Classic: **Sanger sequencing** - Nobel Prize in Chemistry (1980)
- ▶ Second generation/Next-generation(NGS):
 - SOLiD (Thermo Fisher)
 - Ion Torrent (Thermo Fisher)
 - 454 (Roche)
 - HiSeq/MiSeq/... (**Illumina**)
- ▶ Third generation:
 - PACBIO RS (PacBio)
 - MinION**/PromethION/GridION (Oxford Nanopore)
 - ... A lot more emerging technologies/platforms

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

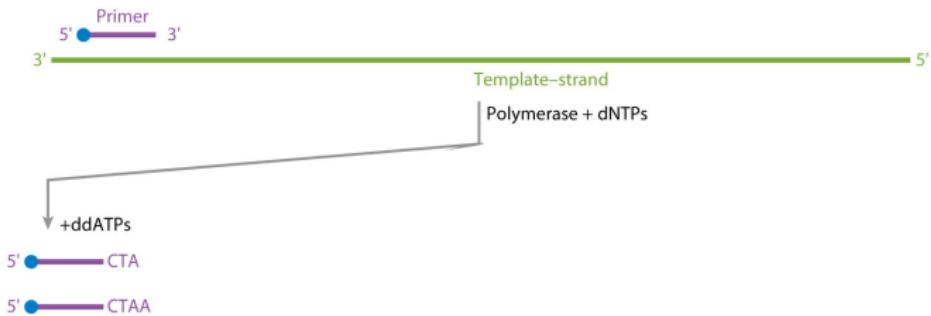
What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Sanger sequencing: 1977



01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

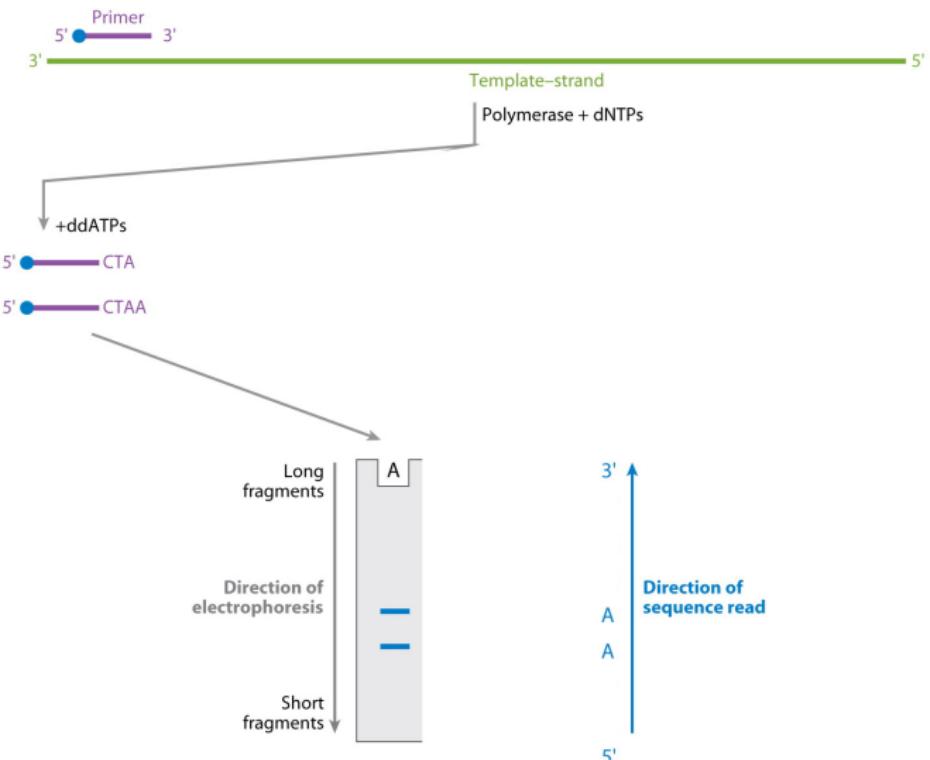
What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Sanger sequencing: 1977



01a-Introduction

- Teaching staff

- Logistics

- Computational Biology

- Brief History of evolution

- Course overview

01b-Sequencing techniques

- About sequences

- What are sequences?

- How can we obtain the sequences?

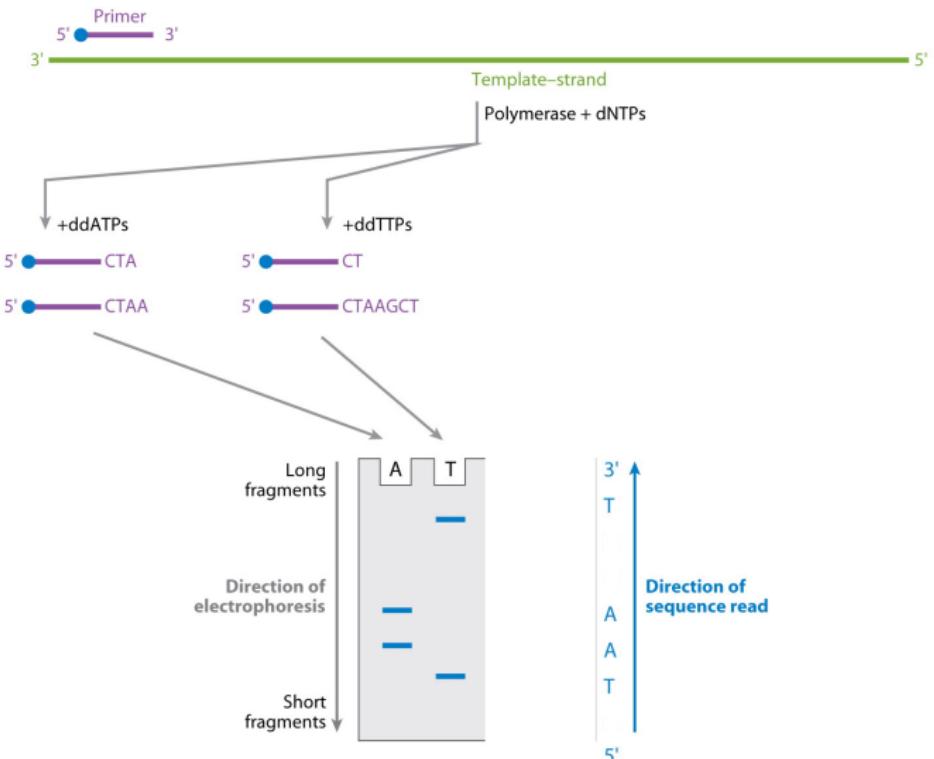
- Sequencing methods

References

Mardis ER. 2013.

Annu. Rev. Anal. Chem. 6:287–303

Sanger sequencing: 1977



01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

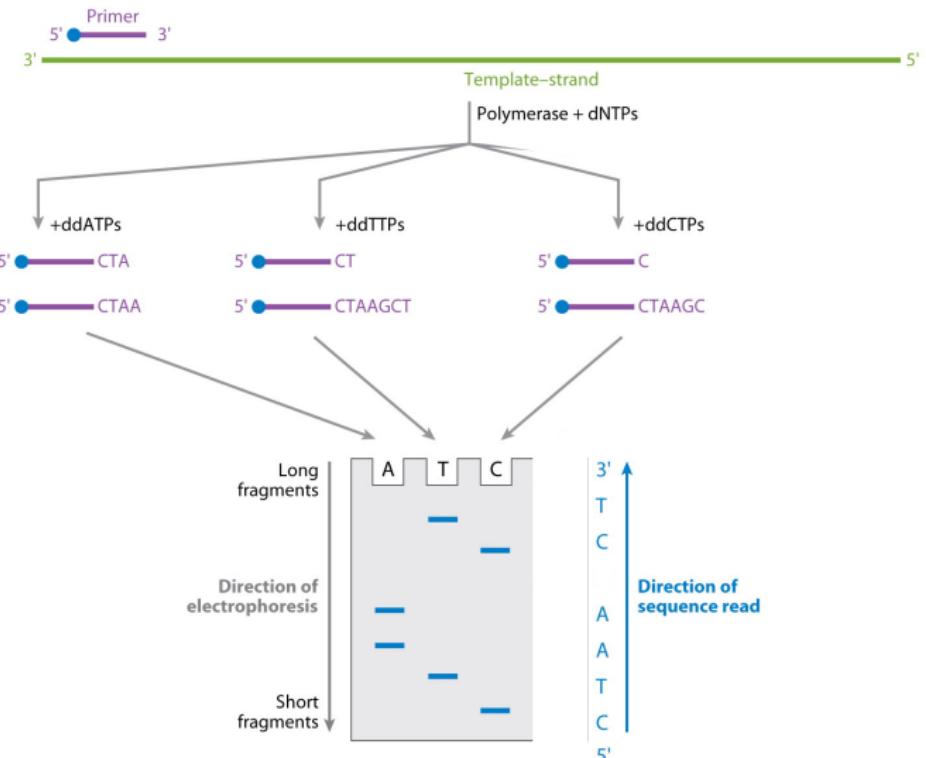
What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Sanger sequencing: 1977



01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

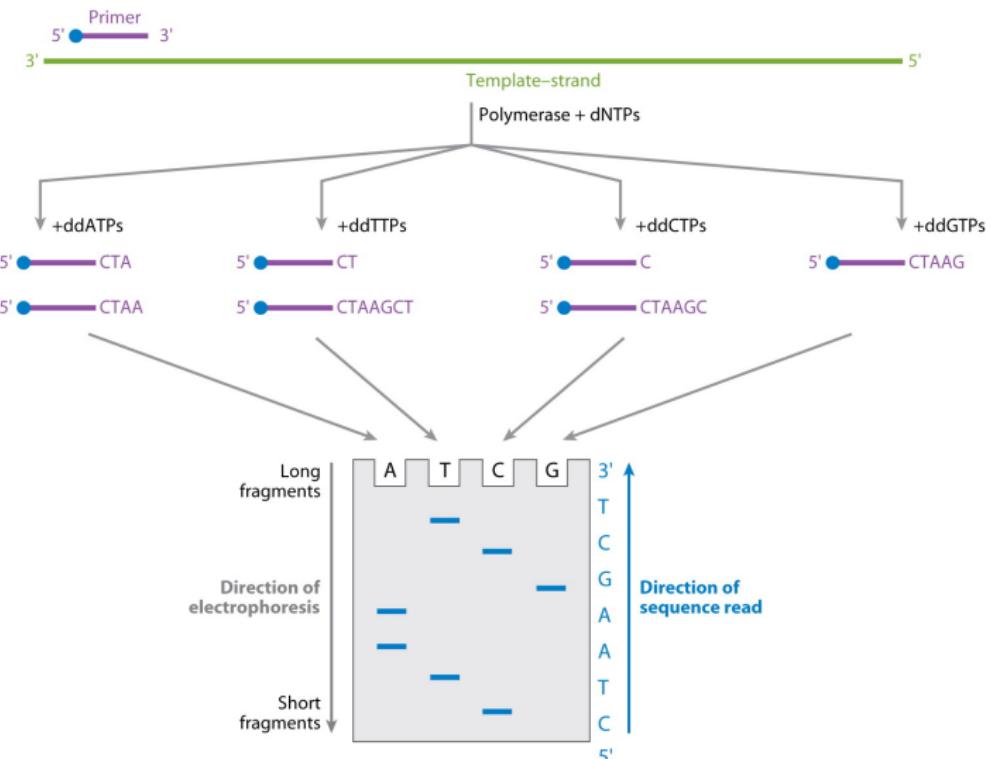
References



Mardis ER. 2013.

Annu. Rev. Anal. Chem. 6:287–303

Sanger sequencing: 1977



01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

Next-generation sequencing (2nd): Illumina

first NGS method developed in 2005, NGS method of the year 2007 (Nature methods)

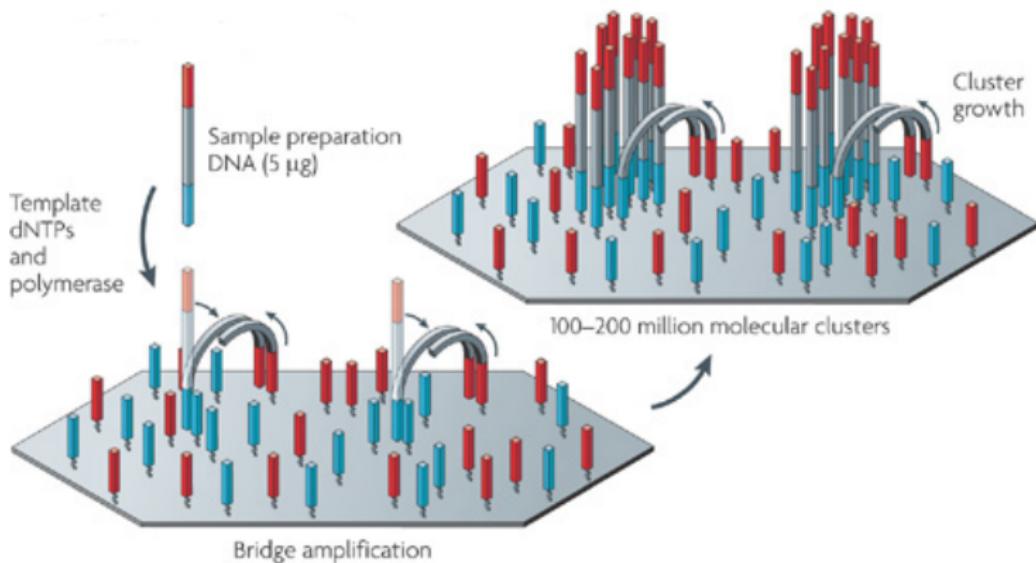


Figure adapted from [Metzker, 2010]

- 01a-Introduction
- Teaching staff
- Logistics
- Computational Biology
- Brief History of evolution
- Course overview

- 01b-Sequencing techniques
 - About sequences
 - What are sequences?
 - How can we obtain the sequences?
 - Sequencing methods

- References

<https://www.youtube.com/watch?v=fCd6B5HRaZ8>

Third-generation: Nanopore

Third generation sequencing: method of the year 2013 (Nature methods.)

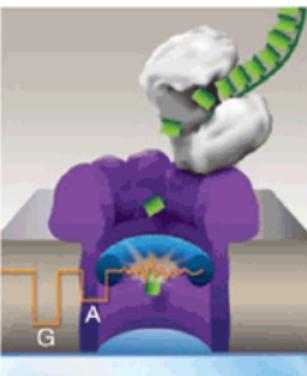
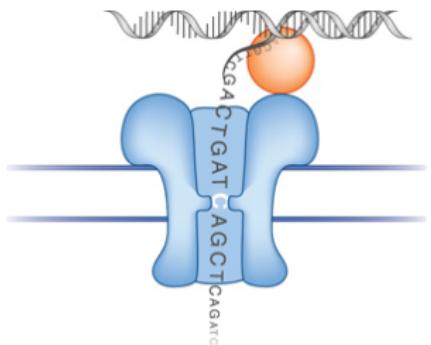


Figure adapted from [Rusk, 2013] and [Munroe and Harris, 2010]

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

<https://www.youtube.com/watch?v=E9-Rm5AoZGw>

Sequencing becomes cheaper and faster

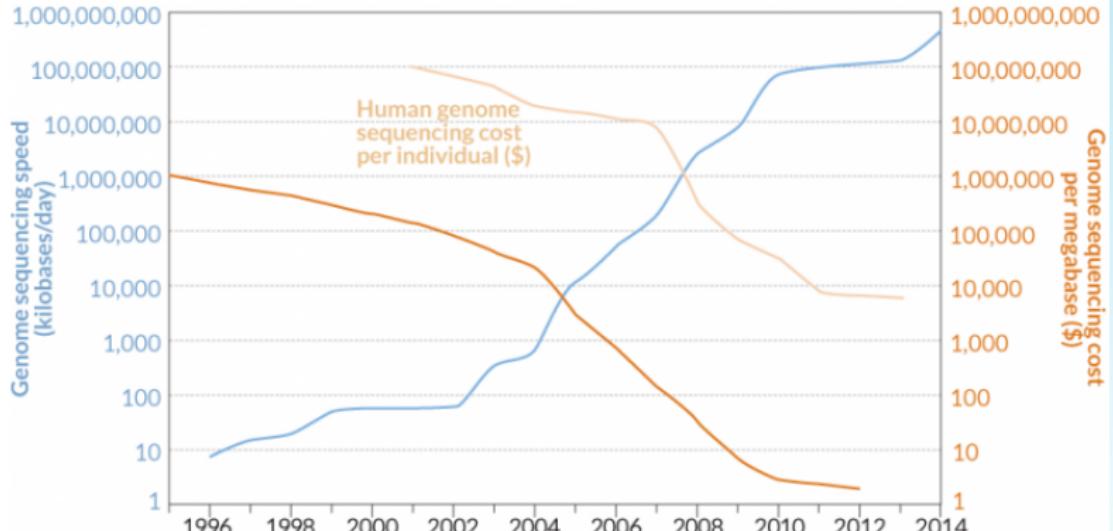


Figure adapted from [ScienceNews.org, 2015]

- 01a-Introduction
- Teaching staff
- Logistics
- Computational Biology
- Brief History of evolution
- Course overview

- 01b-Sequencing techniques
- About sequences
- What are sequences?
- How can we obtain the sequences?
- Sequencing methods
- References

Summary part II

- ▶ overview of sequencing techniques
 - ▶ three generations of sequencing machines
 - ▶ each generation has a similar underlying mechanism

01a-Introduction

- Teaching staff
- Logistics
- Computational Biology
- Brief History of evolution
- Course overview

01b-Sequencing techniques

- About sequences
- What are sequences?
- How can we obtain the sequences?
- Sequencing methods
- References

Further Questions

- ② Name one weakness and one strength for each of the different sequencing methods.
- ② With which of the methods presented could you conceivably sequence the genome of a particular cell?
- ② Imagine you want to perform a paternity test. How would you go about testing whether the potential father is really the father? Could you think of reasons to find a false-negative relatedness?

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References

References |

- adtbio.com (2015). Central dogma of biology.
- Anthro.palomar.edu (2015). Cell to DNA.
- Felsenstein, J. (2004). *Inferring phylogenies*. Sinauer associates Sunderland.
- Fitch, W. M. and Margoliash, E. (1967). Construction of phylogenetic trees. *Science*, 155(3760):279–284.
- ghr.nlm.nih.gov (2015). DNA helix.
- Grenfell, B. T., Pybus, O. G., Gog, J. R., Wood, J. L., Daly, J. M., Mumford, J. A., and Holmes, E. C. (2004). Unifying the epidemiological and evolutionary dynamics of pathogens. *science*, 303(5656):327–332.
- Huxley, J. et al. (1942). Evolution. the modern synthesis. *Evolution. The Modern Synthesis*.
- Metzker, M. L. (2010). Sequencing technologies—the next generation. *Nature reviews genetics*, 11(1):31–46.
- Munroe, D. J. and Harris, T. J. (2010). Third-generation sequencing fireworks at marco island. *Nature biotechnology*, 28(5):426–428.
- Paabo, S. (2003). The mosaic that is our genome. *Nature*, 421(6921):409–12.
- Rusk, N. (2013). Disruptive nanopores. *Nature Methods*, 10(1):35–35.
- ScienceNews.org (2015). Sequencing prices.
- Watson, J. D. and Crick, F. H. C. (1953). Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature*, 171(4356):737–738.
- www.ebi.ac.uk. EBI online training on GWAS.

01a-Introduction

Teaching staff

Logistics

Computational Biology

Brief History of evolution

Course overview

01b-Sequencing techniques

About sequences

What are sequences?

How can we obtain the sequences?

Sequencing methods

References