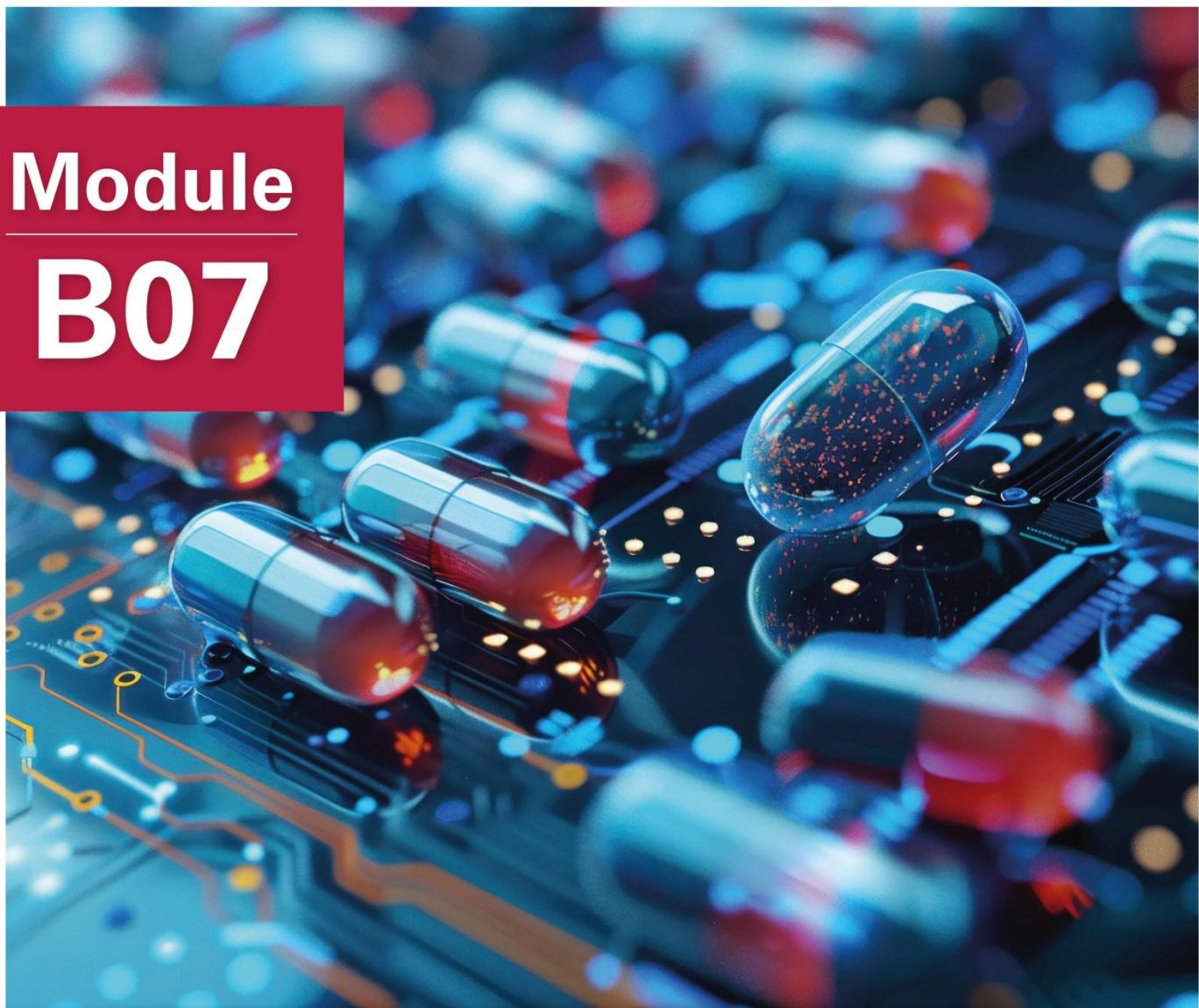


Module **B07**



Bioinformatics Foundational Course

Phylogenetics

NGS Academy for the Africa CDC

Module B07

Phylogenetics

 [back to the table of modules](#)

Module last updated:

December 2024

Suggested or approximate number of sessions	4
Suggested or approximate total learning time	8-10 hours
Target audience	Wet laboratory personnel and bioinformaticians
Delivery format	Lectures, videos, practicals
Level of the module	Introductory to intermediate



Contributors

John Juma, Yiqun Li, Nicky Mulder and Verena Ras.



Suggested prerequisite module(s)

- [Module B01. Introduction to Unix/Linux, Command Line, and Shell Scripting](#)
- [Module B05. Introduction to R Programming for Data Analysis](#) (if you plan to use R for phylogenetics)



Module description

This module provides a detailed look at phylogenetics theory, traits, phylogenies, evolutionary models and divergence times, sequence alignment, building trees using different approaches and different workflows and tools. In this module, participants will specifically cover the following topics and/or concepts:

1. Theory of evolution and trees
 - Historical perspective on the development of the theory of evolution
 - Mechanisms of molecular evolution and speciation
 - Models of DNA and amino acid evolution
 - Theory of phylogenies from biological data

- 
- Evolutionary distances between sequences
 - Rooted versus non-rooted trees
 - Complexity of the space of phylogenetic models, heuristics to map it.
2. Multiple sequence alignment and tree building
- Selecting sequences (getting data from existing databases (see B10), building the final datasets
 - Multiple sequence alignment, correcting/refining alignment
 - Tree building methods:
 - distance matrix methods, Parsimony approaches.
 - Tree building with maximum likelihood Bayesian viewpoint. Branch support measures.
 - Determining temporal signal in dated sequences (Root to tip plot) using TempEST or in R (regression)
3. Phylogenetic tree visualization and Interpretation
- Time trees (e.g. using BEAST or TreeTime)
 - Genomic epidemiology workflows -integrating metadata and annotations (e.g. phylogeography)
 - Visualizing trees with metadata (e.g. using MicroReact, SPREAD (<https://spreadviz.org/>))
 - Understanding and manipulating tree output formats
 - Interpreting a phylogenetic tree
 - Identifying the lineages of the study sequences



Module learning outcomes

On completion of this module, participants will have a basic knowledge of, or will be able to:

- Understand the mathematical modeling of evolution of quantitative or qualitative biological traits
- Describe Markovian processes, and probability distribution of the time between two mutation events, and between two speciation events
- Extract data and perform a multiple sequence alignment
- Effectively build phylogenies from biological sequences using different software, and set the parameters accurately
- Analyze and assess the quality of phylogenetic trees
- Interpret phylogenetic trees



Module assessments

Module practical: Practical available on the [ASLM platform](#)

Module quiz: Assessment questions available on the [ASLM platform](#)



Module resources

- Inferring Phylogenies, by Joseph Felsenstein, published by Sinauer Associates, U.S.A., 2004. ISBN 0878931775, 978-0-878-93177-4.
- Statistical Methods in Molecular Evolution, by Rasmus Nielsen (Ed.) et al., Springer, 2005. ISBN 978-0-387-27733-2.
- Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids, by Durbin, Eddy, Krogh and Mitchison, Cambridge University Press, 1998. ISBN-10: 0521629713 ; ISBN-13: 978-0-521-62971-3
- [H3ABioNet | Webpage - Evolution and Phylogenetics](#)
- [Beast Community | Webpage - BEAST X](#)



Acknowledgements

We would like to thank the following individuals, in alphabetical order of last name, for their valuable time and effort spent in designing (i.e., drafting, reviewing, and refining) this module: **John Juma, Yiqun Li, Nicky Mulder and Verena Ras**.

Furthermore, we would like to thank the following institutions, societies, journals and individuals from whom we sourced open-access resources, used in this module:

Beast Community, H3ABioNet | Pan African Bioinformatics Network for the Human Heredity and Health in Africa (H3Africa) consortium; Richard Durbin, Sean Eddy, Joseph Felsenstein, Anders Krogh, Graeme Mitchison, Rasmus Nielsen.