

# Computational Methods in Evolutionary Biology (IN5088)

---

Titel	Rechnergestützte Methoden in der evolutionären Biologie	
Typ	Vorlesung mit Übung	
Credits	8 ECTS	
Lehrform/SWS	4V + 3Ü	
Sprache	Englisch	
Modulniveau	Master	
Arbeitsaufwand	Präsenzstunden	105 Stunden
	Eigenstudium	135 Stunden
	Gesamtaufwand	240 Stunden
Angestrebte Lernergebnisse	<p>Die Studierenden sind in der Lage, statistische und theoretische Modelle (wie z.B. stochastische Modelle für Evolutionsprozesse und Genealogien, molekulare Uhren, Brownsche Bewegung) geeignet auszuwählen und einzusetzen. Zudem sind sie in der Lage, bioinformatische und statistische Methoden (wie z.B. Maximum Likelihood- und Bayessche Methoden, Markov Chain Monte Carlo (MCMC)) bzw. Programmpakete für die Analyse populationsgenetischer und phylogenetischer Daten (wie z.B. Populationen, Wachstum, Migration) anzuwenden sowie für neue Anwendungen zu entwickeln. Des Weiteren sind sie in der Lage, die erzielten Ergebnisse der Analysen sicher zu interpretieren und kritisch zu überprüfen.</p>	
Intended Learning Outcomes	<p>Students are able to select and to apply statistical and theoretical models (such as stochastic models for evolutionary processes and genealogies, molecular clocks, Brownian motion) as well as to select and apply bioinformatical and statistical methods (such as Maximum Likelihood and Bayesian methods, Markov Chain Monte Carlo (MCMC)) and software packages for the analysis of population genetic and phylogenetic data (such as population, growth, migration). They are also able to develop new algorithms for new problems. Students are also able to interpret and critically judge the results of such analyses.</p>	
Inhalt	<p>Das Modul umfasst rechenintensive Methoden zur Analyse populationsgenetischer Daten (z.B. in Hinblick auf Populationsstruktur, Wachstum und</p>	

	<p>Migrationsraten) und für phylogenetische Analysen. Behandelt werden auch die zu Grunde liegenden stochastischen Evolutionsmodelle und statistische Prinzipien, insbesondere Maximum-Likelihood und Bayessche Methoden, sowie numerische Ansätze wie Importance-Sampling, Markov Chain Monte Carlo (MCMC) und Approximate Bayesian Computation (ABC). Es geht auch darum, welche phylogenetischen Probleme effizient lösbar sind, und welche NP-hart, so dass nur approximative Lösungen verfügbar sind. Behandelt werden auch fortgeschrittene Aspekte der Analyse populationsgenetischer Daten, wie etwa neuere Varianten der ABC-Methodik, anzestrale Rekombinationsgraphen, anzestrale Selektionsgraphen und/oder spezielle Methoden für die Analyse genomweiter Sequenzdaten.</p>
<p>Contents</p>	<p>Contents are computational methods for the analysis of population genetic data (e.g., to infer population structure, growth and migration rates) and for phylogenetic analyses. Covered are also the underlying probabilistic evolution models and statistical principles, in particular Maximum-Likelihood and Bayesian Methods, and computational approaches like importance sampling, Markov Chain Monte Carlo (MCMC) and Approximate Bayesian Computation. We discuss which phylogenetic problems can be solved efficiently and which problems are NP complete, such that only approximate solutions are available. In-depth we treat state-of-the art data analysis methods for special problems in population genetics as for example novel variants of Approximate Bayesian Computation (ABC), Approximations of the Ancestral Recombination Graph, the Ancestral Selection Graph and/or novel methods for analyzing genome-wide sequence data.</p>
<p>Prüfung</p>	<p>Prüfungsleistung (benotet): -Klausur: 120 min</p> <p>Wiederholungsklausur zu Ende des Semesters. Details werden zu Beginn des Moduls bekannt gegeben.</p> <p>In der Klausur weisen die Studierenden nach, inwieweit sie die vorgestellten Modelle, Methoden und Algorithmen verstanden haben, komprimiert wiedergeben und anwenden sowie auf verwandte Problemstellungen übertragen können. In der Klausur werden 7 bis 9 Aufgaben gestellt, die eine eigenständige Anwendung der Modelle, Methoden und Algorithmen aus</p>

	<p>der Vorlesung erfordern (wie z.B. Bayessche Methoden einsetzen, Methoden zur Rekonstruktion von Phylogenien anwenden, stochastische Modelle der Populationsgenetik analysieren, Datenanalyseergebnisse interpretieren).</p>
Examination	<p>Examination requirements (graded): - written exam: 120 min</p> <p>A makeup exam will be offered at the end of the semester; details will be announced at the beginning of the module.</p> <p>Within the written exam, students demonstrate that they understand the presented models, methods, and algorithms, that they can reproduce and apply them as well as that they can transfer and extend models, methods, and algorithms to similar problems. The written exam consist of 7 to 9 assignments, which require independent application of models, methods, and algorithms presented in the lecture (e.g., application of Bayesian Methods, application of phylogeny reconstruction methods, analyses of stochastic population genetic models, interpretation of results of data analyses).</p>
Literatur/Literature	<p>J. Felsenstein (2004) Inferring Phylogenies. Sinauer Z. Yang (2006) Computational Molecular Evolution. Oxford University Press R. Nielsen, [Ed.] (2005) Statistical Methods in Molecular Evolution. Springer R. Durbin, S. Eddy, A. Krogh, G. Mitchison (1998) Biological Sequence Analysis. Cambridge University Press W. Ewens, G. Grant (2005) Statistical Methods in Bioinformatics. Springer Wakeley, J. 2008. Coalescent Theory: An Introduction. Roberts &amp; Company Publishers</p>
Medienformen	<p>Beamer-Präsentation, Tafelpräsentation, Handout, Arbeit am Computer</p>
Media	<p>slides show, blackboard presentation, handouts, computer lab</p>
Lehr- und Lernmethode	<p>Lösen von Übungsaufgaben, Präsentation und Diskussion der Lösungen, Erproben von Software am Computer mit echten und simulierten Daten. Vorlesung, Tutorübung, Rechnerübung, Aufgaben zum Selbststudium. Das Modul besteht aus einer Vorlesung und Übungen in kleinen Gruppen (als Tutor- und Rechnerübungen).</p>

	<p>In den Aufgaben zum Selbststudium, die freiwillig abzugeben sind, analysieren die Studierenden die in der Vorlesung vorgestellten Methoden und Modelle. Mit echten Daten und Simulationsstudien erproben die Studierenden in den Rechnerübungen unter welchen Bedingungen verschiedene Software-Pakete (z.B. Hudson's MS, LAMARC, GENETREE, Ima2, Phylip, RAxML, MrBayes and BEAST) sinnvolle Ergebnisse liefern. Durch das Lösen von Aufgaben, die sich auf die behandelten theoretischen Konzepte beziehen und auch mathematische Herleitungen beinhalten können, vertiefen die Studierenden ihr Verständnis für die Inhalte der Vorlesung. In den Tutorübungen werden mögliche Lösungsansätze der Aufgaben zum Selbststudium diskutiert.</p>
Teaching and Learning Methods	<p>Solving theoretical exercises, presentation and discussion of results of exercises, exploration of computer software, computer simulation studies. Lecture, tutorial, lab course, assignments for individual study.</p> <p>Within the assignments (the submission is optional) students analyze the methods presented in the corresponding lectures. In the lab courses they will also try out various software packages (e.g., Hudson's MS, LAMARC, GENETREE, Ima2, Phylip, RAxML, MrBayes and BEAST) and explore by computer simulation studies under which circumstances they are appropriate. By solving tasks that relate to the theoretical concepts and may include mathematical derivations the students deepen their understanding of the contents of the lecture. Within the tutorials possible approaches for solutions of the assignments will be discussed.</p>
Turnus	Wintersemester
Modulverantwortlicher	Prof. Dr. Dirk Metzler
Dozenten	Prof. Dr. Dirk Metzler