

Praktikum Genomorientierte Bioinformatik (IN5027)

Title	Lab Course Genome-Oriented Bioinformatics	
Typ	Praktikum	
Credits	12	
Lehrform/SWS	10P	
Sprache	Deutsch oder Englisch	
Modulniveau	Bachelor	
Arbeitsaufwand	Präsenzstunden	150 Stunden
	Eigenstudium	210 Stunden
	Gesamtaufwand	360 Stunden
Angestrebte Lernergebnisse	Nach der Teilnahme am Modul sind die Studierenden in der Lage, sich unter Anleitung systematisch im Team in ein klar umrissenes Forschungsprojekt einzuarbeiten, die vorgegebenen Teilaufgaben zeitlich zu planen und umzusetzen (ggf. Anpassungen am Zeitplan vorzunehmen) und die erzielten Ergebnisse zu bewerten und zu diskutieren sowie die Resultate zu präsentieren. Die Studierenden können unter Anleitung in Gruppen die Thematik und den biologischen bzw. biotechnologischen Hintergrund eines Forschungsprojekts (z.B. aus dem Bereich der Proteinstruktur- und Funktionsvorhersage, Genomics, Epigenomics, Transkriptomik, Proteomik, Analyse von NGS-Daten, Vorhersage und Analyse regulatorischer Netzwerke, genetisch bedingte Krankheiten) mithilfe eines aktuellen Literaturstudiums und von Vorträgen erarbeiten. Sie können für das konkrete Forschungsprojekt den Zeitplan der vorgegebenen Meilensteine definieren, so dass diese in einem Team von 3-5 Mitgliedern in der vorgegebenen Zeit bearbeitet werden können. Die Studierenden können die für die Lösung benötigten bioinformatischen Methoden (wie z.B. Mapping, differentielle Expression, Splicing Detection, Enrichment Analysis und Pathway Mapping für NGS-Datenanalyse) und existierende Softwaretools bzw. Softwareplattformen anwenden und in der Regel an die gegebenen Anforderungen anpassen bzw. neu entwerfen und darauf aufbauend ein Programmpaket zur Lösung entwickeln. Die Studierenden können die erzielten Ergebnisse evaluieren und mit den vorgegeben Gold-Standards die Ergebnisse daran validieren sowie diese	

	Resultate in einem Vortrag präsentieren und im Rahmen eines wissenschaftlichen Diskurses diskutieren und diese dann im Stile einer wissenschaftlichen Publikation zusammenfassen.
Intended Learning Outcomes	After the successful participation, students are able to familiarize themselves under supervision systematically within a team with a given well-defined research project, to schedule work packages of the research project (including modifications if necessary), and to evaluate and to discuss the results. The students are able to become acquainted under supervision in a team with the biological and biotechnological background of the research project (e.g., from the area of protein structure and function prediction, genomics, epigenomics, transcriptomics, proteomics, analysis of NGS data, prediction and analysis of regulatory networks, genetic diseases) using studies of current literature and presentations. The students are able to specify the timelines of the given milestones such that it can be completed in time with 3 to 5 teammates. The students are able to apply, and usually to adapt or even to redesign (with respect to the actual requirements) bioinformatics methods (e.g., mapping, differential expression, splicing detection, enrichment analysis, pathway mapping for NGS) as well as existing software packages and platforms, and finally to build a software package as a solution for the given research project. The students are able to evaluate the results, to validate these results using given gold standards as well as to present the findings in a scientific talk and to summarize these in a scientific publication.
Inhalt	Das Forschungsprojekt wird jeweils vom ausrichtenden Lehrstuhl vorgegeben. Er orientiert sich an aktuellen Forschungsthemen der Lehrstühle der Bioinformatik (z.B. -omics Datenanalyse, Systembiologie, Analyse und Interpretation biologischer Netzwerke, Analyse von Proteinstruktur- und Proteinfunktionsdaten, genetisch bedingte Krankheiten) und an aktuellen Forschungsfragestellungen der Bioinformatik, die sich ergeben (i) aus neuen biotechnologischen Messtechniken (z.B. DNA-Chips, Sequenzierung menschlicher Genome, Hochdurchsatz-Genom- und -Transkriptom-Sequenzierung (NGS), Long-read Sequenzierung (TGS), Epigenetik und Epigenomik, Metagenomik, Proteomik, Einzelzell-Sequenzierung,

	räumliche Transkriptomik, Gewebe-Karten (ii) aus der Veröffentlichung neuer umfassender Datensätze (wie ENCODE, Epigenomics Roadmap, TCGA, GTEx, Human und Mouse Cell Atlas, Heart Cell Atlas, HuBMAP) und Datenbanken (Swissprot, PDB, Ensembl, GEO, SRA, PRIDE, Proteomics DB, ProteomeXchange) und (iii) aus neuen algorithmischen Verfahren (z.B. zur Analyse von NGS reads, single-cell und spatial transcriptomics Daten), aus neuen Methoden des maschinellen Lernens (Deep Learning, Embedding) und aus neuen KI Methoden (AlphaFold2, Bayes'sche und kausale Modellierung) ergeben. Forschungsprojekte können auch die Vorbereitung und Beteiligung an internationalen Bioinformatikwettbewerben wie CASP und CAFA (Proteinstruktur bzw. -funktionsvorhersage), BioCreative (Text Mining), DREAM (Analyse von Netzwerken, Expression und komplexer Krankheiten) sein.
Contents	The research project will be defined by the respective responsible chair. It is selected from the research areas of the bioinformatics research groups (e.g. -omics data analysis, systems biology, analysis and interpretation of biological networks, analysis of protein structure and function, genetically determined diseases) and from current research problems in bioinformatics, implied by (i) new biotechnological measurements (such as DNA-chips, sequencing of human genomes, high-throughput sequencing techniques for genomes and transcriptomes (NGS), long-read sequencing (TGS), epigenetics and epigenomics, metagenomics, proteomics, single-cell-sequencing, spatial transcriptomics, tissue maps), (ii) by the availability of new comprehensive data sets (such as ENCODE, Epigenomics Roadmap, TCGA, GTEx, human and mouse cell atlas, human heart atlas, HuBMAP) or databases (such as Swissprot, PDB, Ensembl, GEO, SRA, PRIDE, Proteomics DB, ProteomeXchange), and (iii) by new algorithmic methods (e.g. for the analysis of NGS reads, single-cell and spatial transcriptomics data), by new method of machine learning (deep learning, embedding) and by new AI methods (AlphaFold2, Bayesian and causal modeling). Research projects could also be the preparation and participation at international bioinformatics competitions such as CASP and CAFA (protein

	structure and function prediction, respectively), BioCreative (text mining), DREAM (network, expression and complex disease analysis).
Prüfung	Die Prüfung ist eine benotete Laborleistung und besteht aus mehreren Prüfungselementen (Präsentationen, Hausaufgaben, zu erstellende Programme, Projekte und Ausarbeitungen). Dadurch werden verschiedene Lernergebnisse überprüft: Durch Präsentationen (je nach Thema etwa 2-4 Kurzvorträge à 20-30 Minuten pro Studierenden) wird nachgewiesen, dass die Studierenden erarbeitetes Wissen (aus dem Literaturstudium und im Projekt gewonnene Erkenntnisse) und Ergebnisse in Form von Seminarvorträgen darstellen können. Durch Hausaufgaben, den Entwurf und die Implementierung von Programmen sowie die Realisierung von Projekten weisen die Studierenden nach, dass sie erworbenes Wissen (wie z.B. aktuelle bioinformatische Methoden) praktisch anwenden und umsetzen können. In den abschließenden Ausarbeitungen weisen die Teilnehmer nach, dass sie ihre Ergebnisse wissenschaftlich beschreiben und ihre Arbeit in Bezug auf den aktuellen Stand der Wissenschaft bewerten können. Die Gesamtnote ergibt sich aus der Kombination der Prüfungselemente (Präsentationen, Hausaufgaben, zu erstellende Programme, zu realisierende Projekte und Ausarbeitungen). Die genaue Anzahl der Vorträge, der Hausaufgaben, Programme, Projekte, Ausarbeitungen und die Gewichtung der Teile für die Gesamtnote wird jeweils zu Beginn des Moduls bekannt gegeben. Die Prüfung kann im Folgejahr wiederholt werden.
Examination	The exam is a graded lab exam consisting of several examination elements (presentations, homework assignments, programs, projects, and reports). Thereby different learning outcomes of the lab course genome oriented bioinformatics will be assessed: Learned knowledge and results (acquired in the literature study and in the project progress) are presented in the form of seminar talks (depending on the main topic approximately 2-4 talks lasting 20-30 minutes each per student). The application and transfer of learned knowledge (e.g., state-of-the-art bioinformatics methods) are assessed via homework assignments and via the

	design and implementation of programs and the realization of projects. In the final reports the participants describe their results according to scientific standards and discuss their work in the context of the current state-of-the art. The final grade of the module is determined from the combination of the examination elements (presentations, homework assignments, programs, projects, and reports). The specific number of talks, homework assignments, programs, projects, and reports as well as the weighting of the respective parts is fixed and communicated at the beginning of the module. The exam can be retaken in subsequent year.
Literatur/Literature	Aktuelle Fachartikel zum Thema
Medienformen	Folienpräsentation, Tafelanschrieb, Computerlabor
Media	slide show, blackboard, computer lab
Lehr- und Lernmethode	Vorträge, Seminar, Rechnerpraktikum, Aufgaben zum Selbststudium, Gruppenarbeit, Projekt Biologisches und biotechnologisches Hintergrundwissen sowie grundlegende Modelle und Methoden, die für die Realisierung des Projekts benötigt werden, werden durch einführende Vorträge der Dozenten und kurze Spezialvorträge der Studierenden vermittelt. Im Seminar werden diese Spezialvorträge sowie der Fortschritt der Projekte diskutiert. Das Projekt selbst wird dabei im Team am Rechner implementiert, wobei die Lösung der Aufgaben zum Selbststudium die Implementierung erleichtert.
Teaching and Learning Methods	Talks, seminar, computer lab course, assignments for individual study, team work, project Biological and biotechnological background knowledge as well as fundamental models and methods used for the realization of the project will be presented through introductory talks given by the instructors as well as through specific talks given by the students. Within the seminar the topics of the specific talks as well as the progress of the projects will be discussed. The project will be implemented in a team using the computer, for which the solution of the assignments for individual study supports the implementation.

Turnus	Wintersemester
Modulverantwortlicher	Prof. Dr. Ralf Zimmer
Dozenten	Prof. Dr. Ralf Zimmer