

# Perlen der Bioinformatik: ENCODE (IN5096)

---

Title	Bioinformatics Pearls: ENCODE	
Typ	Vorlesung mit Übungen	
Credits	9	
Lehrform/SWS	4V + 2Ü	
Sprache	English	
Modulniveau	Master	
Arbeitsaufwand	Präsenzstunden	90 Stunden
	Eigenstudium	180 Stunden
	Gesamtaufwand	270 Stunden
Angestrebte Lernergebnisse	<p>Die Studierenden sind in der Lage, Analysen komplexer Hochdurchsatzexperimente (z.B. ENCODE, TCGA, Epigenomics Roadmap) zu verstehen und zu reproduzieren. Dazu können sie involvierte Originalartikel verstehen und die zur Transkriptom-Analyse verwendeten Methoden und Algorithmen nachvollziehen. Studierende können die erlernten Techniken (wie die Verfahren für die unten aufgeführten einzelnen Themen aus dem ENCODE Projekt) auf neue Hochdurchsatzdatensätze anwenden und für neue Gegebenheiten effizient anpassen.</p>	
Intended Learning Outcomes	<p>Students are able to understand and to reproduce analyses of complex high throughput experiments (e.g. ENCODE, TCGA, Epigenomics Roadmap). For this, students can understand involved original publications and methods and algorithms used for transcriptome analysis. Students can apply the learned techniques (e.g. the methods for the below listed topics from the ENOCODE project) to new high-throughput data sets and can efficiently adapt them to new problem settings.</p>	
Inhalt	<p>Zehn Jahre nach dem Abschluss des menschlichen Genomprojektes, beinhaltet das ENCODE Projekt den nächsten revolutionären Meilenstein für die Genom- und System-Bioinformatik. Ähnliche Durchbrüche für die Krebsforschung stellen TCGA (The Cancer Genome Atlas) und für die Epigenetik die Epigenomics Roadmap dar.</p> <p>In der Nature-Ausgabe vom 6. September 2012, feiert das ENCODE-Projekt seine Fertigstellung mit sechs Artikeln plus Kommentaren und mehr als 20 zusätzlichen wissenschaftlichen Artikeln in anderen</p>	

	<p>Zeitschriften wie Genome Research und Genome Biology.</p> <p>Die Vorlesung präsentiert das ENCODE Projekt und die dazugehörigen Publikationen. Sie liefert eine gründliche Einführung in die Grundlagen, aktuellen Ergebnisse und zukünftige Perspektiven von ENCODE für die Bioinformatik. Die Studierenden werden tiefgehende praktische Erfahrungen sammeln bezüglich der ENCODE Daten, ihrer Analyse und Verwendung. Die Techniken und Datentypen, die von ENCODE verwendet wurden, werden weitere Entwicklungen in der molekularen und Genombiologie bewirken, sowie der Bioinformatik und Systembiologie. ENCODE wird dabei die Basis liefern und eine der wichtigsten Schlüsselressourcen eines großen Teils der bioinformatischen Forschung in den nächsten Jahren.</p> <p>Die folgenden Themen werden dabei abgedeckt:</p> <ol style="list-style-type: none"> <li>1. Transkriptionsfaktormotive</li> <li>2. Chromatin Muster an Transkriptionsfaktorbindestellen</li> <li>3. Charakterisierung von intergenischen Bereichen und Gendefinition</li> <li>4. RNA und Chromatin-Modifikationsmuster um Promotoren herum</li> <li>5. Epigenetische Regulierung von RNA-Prozessierung</li> <li>6. Charakterisierung von nicht-kodierender RNA</li> <li>7. DNA Methylierung</li> <li>8. Enhancer-Bestimmung und Charakterisierung</li> <li>9. Drei-dimensionale Verbindungen im Genom</li> <li>10. Charakterisierung von Netzwerktopologie</li> <li>11. Machine Learning Ansätze in der Genomik</li> <li>12. Einfluss von funktioneller Information auf das Verständnis von Variation</li> <li>13. Einfluss von evolutionärer Selektion auf funktionelle Bereiche</li> </ol>
<p>Contents</p>	<p>Ten years after the completion of the human genome project, the ENCODE project marks the next revolutionary milestone for genome and systems bioinformatics. Similar breakthroughs for cancer research are TCGA (The Cancer Genome Atlas) and for Epigenetics the Roadmap.</p> <p>In the September 6, 2012, issue of Nature, the ENCODE project celebrates its completion with six articles plus</p>

	<p>commentaries and more than 20 additional scientific articles in other journals such as Genome Research and Genome Biology.</p> <p>The lecture series will present the ENCODE project and the accompanying publications. It will provide a thorough introduction into the foundations, the current results and the future prospects of ENCODE for Bioinformatics. Students will also get an in-depth practical hands-on experience with ENCODE data, its analysis and use. The techniques and data types employed in ENCODE will spur the further development of molecular and genome biology, bioinformatics and systems biology. ENCODE is expected to be the basis and one of the most important key resources for a large part of bioinformatics research for the years to come.</p> <p>The following topics will be addressed:</p> <ol style="list-style-type: none"> <li>1. Transcription factor motifs</li> <li>2. Chromatin patterns at transcription factor binding sites</li> <li>3. Characterization of intergenic regions and gene definition</li> <li>4. RNA and chromatin modification patterns around promoters</li> <li>5. Epigenetic regulation of RNA processing</li> <li>6. Non-coding RNA characterization</li> <li>7. DNA methylation</li> <li>8. Enhancer discovery and characterization</li> <li>9. Three-dimensional connections across the genome</li> <li>10. Characterization of network topology</li> <li>11. Machine learning approaches to genomics</li> <li>12. Impact of functional information on understanding variation</li> <li>13. Impact of evolutionary selection on functional regions</li> </ol>
Prüfung	<p>Prüfungsleistung (benotet): -Klausur: 120 min</p> <p>Wiederholungsklausur zu Ende des Semesters. Details werden zu Beginn des Moduls bekannt gegeben.</p> <p>Mit der Klausur wird geprüft, inwieweit die Studierenden Konzepte und Algorithmen zur Analyse von Hochdurchsatzdaten und Transkriptomen</p>

	<p>verstehen und komprimiert in begrenzter Zeit wiedergeben können. Anhand konkreter Problemstellungen wird geprüft, ob Studierende die Algorithmen und Lösungen (für die oben genannten 13 Probleme aus dem ENCODE Projekt) an verwandte Problemstellungen anpassen können. Konkret werden in der Klausur 4-7 Aufgaben bearbeitet, die eine eigenständige Anwendung der Konzepte, Modelle und Methoden aus der Vorlesung zur Lösung einer anspruchsvollen Problemstellung erfordern (wie z.B. Vorhersage von TF Bindestellen, Identifikation von Netzwerkmustern, Korrelation von Chromatin mit Epigenetischen Markern und Expressionsdaten) oder die Analyse von bestimmten ENCODE-Daten (wie z.B. Expressionsdaten, Chromatin Modifikationen, epigenetische Datensätze).</p>
Exam	<p>Exam (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 course.</p> <p>Within the written exam, students demonstrate that they understand concepts and algorithms for the analysis of high-throughput data and transcriptomes and they can reproduce it in condensed form within in limited time. For specific problems, the students demonstrate they are able to adapt the algorithms and solutions (for the 13 above mentioned problems of the ENCODE project) to similar problems. The written exam contains 4 to 7 assignments, which require independent application of concepts, models and methods presented in the lecture to solve demanding problems (e.g., prediction of TF binding sites, identification of network motifs, correlation of chromatin with epigenetic markers and expression data) or the analysis of certain ENCODE data (e.g., expression data, chromatin modification, epigenetic data).</p>
Literatur	Nature Sep 6, 2012 (ENOCDE issue, Vol 489, No 7414, p. 1-170)
Literature	Nature Sep 6, 2012 (ENOCDE issue, Vol 489, No 7414, p. 1-170)
Medienformen	Folienpräsentation, Tafelanschrieb
Media	Slide show, blackboard
Lehr- und	Vorlesung, Tutorübung, Aufgaben zum Selbststudium.

Lernmethode	<p>Das Modul besteht aus einer Vorlesung sowie Übungen in kleineren Gruppen.</p> <p>In den Hausaufgaben, die freiwillig abzugeben sind, wird das Verständnis der Konzepte, Modelle und Methoden (die in der Vorlesung vorgestellt werden) anhand konkreter Daten (in der Regel aus ENCODE) und Beispielen vertieft. Die Studierenden übertragen Konzepte und Modelle und analysieren Modelle und Daten mithilfe der in der Vorlesung vorgestellten Konzepte, Modelle und Methoden und diskutieren die Ergebnisse. In den Hausaufgaben werden selbständig anspruchsvolle Übungsaufgaben bearbeitet, die ähnlich wie die Klausuraufgaben sind (siehe oben) und deshalb zur Vorbereitung darauf dienen.</p> <p>In den Übungen werden mögliche Lösungsansätze der Aufgaben zum Selbststudium diskutiert.</p>
Teaching and Learning Methods	<p>Lecture, tutorial, assignments for individual study. The module consist of a lecture and in addition exercises in small groups.</p> <p>Within the assignments (submission is optional), models and methods (presented in the lecture) will be applied to real data and examples to deepen the understanding. The students transfer concepts and models, analyze models and data (based on the presented concepts, models, and methods), and discuss the results. The assignments consist of demanding problems similar to the assignments in the written exam (for details see above) and serve as a preparation for the written exam. Within the tutorials possible approaches for the solutions of the assignments will be discussed.</p>
Turnus	unregelmäßig
Modulverantwortlicher	Prof. Dr. Ralf Zimmer
Dozenten	Prof. Dr. Ralf Zimmer Prof. Dr. Caroline Friedel