

DENOMINAZIONE INSEGNAMENTO / ATTIVITÀ FORMATIVA

METODOLOGIE PER L'ANALISI DI DATI GENOMICI

METHODS FOR THE ANALYSIS OF GENOMIC DATA

INFORMAZIONI INSEGNAMENTO / ATTIVITÀ FORMATIVA

A.A.	2026-2027	CdS	BIOLOGIA AMBIENTALE
Codice	8068330	Canale	Unico
CFU	3+3	Lingua	Italiano

DOCENTE RESPONSABILE DELL'INSEGNAMENTO / ATTIVITÀ FORMATIVA

Paolo Gratton

MODULO DIDATTICO

Claudio Ottoni

Metodologie per l'analisi di dati genomici (3 CFU)

Methods for the analysis of genomic data (3 CFU)

OBIETTIVI FORMATIVI E RISULTATI DI APPRENDIMENTO ATTESI

Italiano	OBIETTIVI FORMATIVI: L'obiettivo dell'insegnamento è quello di acquisire nuove conoscenze e competenze sull'utilizzo delle tecnologie e dei metodi di analisi più comunemente adottati nell'analisi di dati genomici (genomi interi e dati "genome-wide"). Particolare attenzione è rivolta all'uso di tali metodi nella ricostruzione di processi evolutivi delle popolazioni umane e animali, acquisendo capacità critiche per comprenderli e interpretarli. CONOSCENZA E CAPACITÀ DI COMPRENSIONE: Apprendimento delle tecniche usate per sequenziare genomi interi e dei metodi più comunemente usati in genomica di popolazioni e in bioinformatica per analizzare i dati di sequenziamento high-throughput (NGS). CAPACITÀ DI APPLICARE CONOSCENZA E COMPRENSIONE: Acquisizione di capacità per affrontare un problema sperimentale nell'analisi dei dati genomici con gli adeguati strumenti cognitivi. Acquisizione di competenze applicative di tipo metodologico, tecnologico e strumentale nello studio dei dati genomici. AUTONOMIA DI GIUDIZIO: Acquisizione di competenze per sviluppare linee di ricerca nell'ambito della genomica per lo studio dell'evoluzione e interpretare in modo consapevole i dati sperimentali; essere in grado
----------	--

	di eseguire ricerche bibliografiche e di selezionare criticamente i materiali di interesse, in particolare sul web. ABILITÀ COMUNICATIVE: Essere in grado di organizzare, presentare in modo efficace anche graficamente e trasmettere le proprie conoscenze o i risultati della più recenti scoperte nell'ambito della genomica per lo studio dell'evoluzione sia in forma scritta che orale, anche in lingua inglese. Queste capacità sono acquisite attraverso interazione costante con gli studenti durante le lezioni frontali, attraverso le risposte a eventuali domande o chiarimenti e la partecipazione a sessioni pratiche di analisi bioinformatica dei dati. CAPACITÀ DI APPRENDIMENTO: Capacità di apprendere e di aggiornarsi in modo autonomo e continuo attingendo a diverse fonti, anche in lingua inglese, con particolare riferimento alla consultazione critica di materiale bibliografico anche di livello avanzato, di banche dati e di altre informazioni in rete.
<i>English</i>	LEARNING OUTCOMES: The course aims to provide students with new knowledge and skills in the use of the technologies and analytical methods most commonly employed in the analysis of genomic data, including whole-genome and genome-wide datasets. Particular emphasis is placed on the application of these methods to the reconstruction of evolutionary processes in human and animal populations, fostering the critical skills necessary for their understanding and interpretation. KNOWLEDGE AND UNDERSTANDING: Students will acquire knowledge of the techniques used for whole-genome sequencing and of the methods most commonly employed in population genomics and bioinformatics to analyse NGS data. APPLYING KNOWLEDGE AND UNDERSTANDING: Students will develop the ability to address experimental problems in genomic data analysis using appropriate conceptual and analytical tools. They will acquire methodological, technological, and practical skills for the analysis and interpretation of genomic datasets. MAKING JUDGEMENTS: Students will acquire the skills necessary to develop research approaches in genomics for the study of evolutionary processes and to interpret experimental data critically and independently. They will also develop the ability to conduct literature searches and to critically evaluate and select relevant scientific sources, particularly those available online. COMMUNICATION SKILLS: Students will be able to organize, present, and effectively communicate their knowledge and the results of recent advances in evolutionary genomics, both orally and in writing, including

	in English. These skills will be developed through continuous interaction during lectures, discussion sessions, responses to questions and requests for clarification, and participation in practical bioinformatics exercises involving genomic data analysis. LEARNING SKILLS: Students will develop the ability to learn independently and to continuously update their knowledge by drawing on a variety of information sources, including English-language resources. Particular emphasis will be placed on the critical use of scientific literature, including advanced publications, specialized databases, and other online resources.
--	---

PREREQUISITI

Italiano	Conoscenze di base di Genetica di Popolazioni e Biologia Evolutiva
<i>English</i>	<i>Basic knowledge of Population Genetics and Evolutionary Biology</i>

PROGRAMMA E CRONOPROGRAMMA

Italiano	LEZIONI FRONTALI (24 h) Il Corso si articola in 24 ore di lezioni frontali (3 CFU), divise in tre parti principali: Prima parte (8 h): Introduzione del corso, definizioni. Costruzione delle librerie genomiche, piattaforme di sequenziamento Illumina, il sequenziamento SBS (Sequencing by Synthesis). Tecniche di arricchimento, sequenziamento shotgun, assemblaggio dei genomi da dati di sequenza (de novo e allineamento a sequenze di riferimento). Seconda parte (10 h): Formati dei file (fasta, fastq, sam, bam). Pannelli di SNPs (formato plink e eigenstrat). Introduzione a Unix, allineamento di fastq a una sequenza di riferimento (è previsto l'uso del computer). Analisi dei files BAM. Terza parte (6 ore): statistiche D e F in genomica e principi di genomica di popolazioni. Utilizzo del notebook Jupyter e analisi di pannelli di SNPs.
<i>English</i>	LECTURES (24 hours) The course consists of 24 hours of lectures (3 ECTS credits), divided into three main sections: Part One (8 hours): Course introduction and key definitions. Construction of genomic libraries, Illumina sequencing platforms, and Sequencing by Synthesis (SBS) technology. Target enrichment techniques, shotgun sequencing, and genome assembly from sequencing data (de novo assembly and alignment to reference genomes). Part Two (10 hours): File formats commonly used in genomics and bioinformatics (FASTA, FASTQ, SAM, and BAM). SNP panels (PLINK and EIGENSTRAT formats). Introduction to the Unix operating system. Alignment of FASTQ files to a reference genome (including hands-on computer-based exercises). Analysis and processing of BAM files.

	Part Three (6 hours): D-statistics and F-statistics in genomics, and fundamental concepts of population genomics. Introduction to the Jupyter Notebook environment and analysis of SNP datasets.
--	---

TESTI ADOTTATI E BIBLIOGRAFIA DI RIFERIMENTO

Italiano	Slide delle lezioni e articoli scientifici recenti forniti dal docente. Tutorial website: https://paleogenomics-tor-vergata.readthedocs.io/en/latest/index.html
English	<i>Articles and other material provided during the lectures.</i> <i>Tutorial website: https://paleogenomics-tor-vergata.readthedocs.io/en/latest/index.html</i>

DESCRIZIONE DELLE MODALITÀ DI SVOLGIMENTO E METODI DIDATTICI ADOTTATI

☒ In presenza

☐ A distanza

Italiano	Lezioni frontali interattive supportate da presentazioni digitali e discussione critica di articoli scientifici chiave. Uso del computer nelle sessioni di analisi dei dati.
English	<i>Interactive classroom lectures supported by digital presentations and critical discussion of key scientific papers. Use of computer in the data analysis sessions.</i>

MODALITÀ DI FREQUENZA

☐ frequenza obbligatoria

☒ frequenza facoltativa

DESCRIZIONE DELLE MODALITÀ DI FREQUENZA

Italiano	La frequenza delle lezioni frontali è caldamente consigliata.
English	<i>Attendance of classroom lectures is highly recommended.</i>

MODALITÀ DI VALUTAZIONE

☐ Prova scritta

☒ Prova orale

☐ Prova di laboratorio

☐ Prova pratica

☐ Valutazione in itinere

☐ Valutazione di progetto

☐ Valutazione di tirocinio

COMPOSIZIONE DELLA COMMISSIONE VALUTATRICE

Claudio Ottoni, Paolo Gratton, Gabriele Scorrano, Cristina Martinez-Labarga

DESCRIZIONE DELLE MODALITÀ E DEI CRITERI DI VERIFICA DELL'APPRENDIMENTO

Italiano	L'esame finale consiste in una prova orale sui temi trattati, tesa a verificare la comprensione profonda dei metodi di sequenziamento NGS e l'analisi dei dati di sequenza. Per superare l'esame occorre riportare un voto non inferiore a 18/30; più in particolare, la votazione è così determinata: 18-21: lo studente ha acquisito i concetti di base e ha una capacità di analisi che emerge solo con l'aiuto del docente. Il modo di esprimersi e linguaggio usato sono complessivamente corretti. 22-25: lo studente ha acquisito in maniera discreta i concetti di base, sa orientarsi tra i vari argomenti trattati e possiede una capacità di analisi autonoma che sa esprimere con un linguaggio corretto. 26-29: lo studente possiede un bagaglio di conoscenze ben strutturato. È in grado di rielaborare in maniera autonoma le conoscenze acquisite con ricchezza di riferimenti che sa esprimere con un linguaggio corretto. 30 e 30 e lode: lo studente possiede un bagaglio di conoscenze completo e approfondito. I riferimenti culturali sono ricchi e aggiornati. Si esprime con brillantezza e proprietà di linguaggio. Non idoneo: importanti carenze e/o inaccurately nella conoscenza e comprensione degli argomenti; limitate capacità di analisi e sintesi, frequenti generalizzazioni e limitate capacità critiche e di giudizio, gli argomenti sono esposti in modo non coerente e con linguaggio inappropriato.
English	<i>The final assessment consists of an oral examination covering the topics addressed during the course and is designed to evaluate students' in-depth understanding of next-generation sequencing (NGS) technologies and genomic data analysis.</i> <i>To pass the examination, students must obtain a grade of at least 18/30. Grades are awarded according to the following criteria:</i> 18–21: <i>demonstrates a basic understanding of the fundamental concepts and analytical skills that emerge only with guidance from the examiner. Scientific terminology and communication are generally correct.</i> 22–25: <i>demonstrates a satisfactory understanding of the core concepts, is able to navigate the topics covered during the course, and shows an autonomous capacity for analysis expressed through appropriate scientific language.</i> 26–29: <i>demonstrates a well-structured and comprehensive body of knowledge. Is able to critically elaborate on the concepts acquired, integrating them with relevant examples and references, and communicates using accurate scientific terminology.</i> 30–30 cum laude: <i>demonstrates an extensive and in-depth mastery of the subject. Knowledge is broad, up-to-date, and critically integrated. Ideas are communicated with exceptional clarity, precision, and scientific rigor.</i> Fail: <i>significant gaps and/or inaccuracies in the knowledge and understanding of the topics covered; limited analytical and synthetic abilities; frequent overgeneralizations; insufficient critical thinking and judgment; inability to present concepts coherently and appropriately using scientific terminology.</i>

