

NAZIV PREDMETA		Bioinformatika				
Kod	PMB704	Godina studija	1.			
Nositelji predmeta	Doc. dr .sc. Željka Trumbić Doc. dr .sc. Matilda Šprung	Bodovna vrijednost (ECTS)	5			
Suradnici	/	Način izvođenja nastave (broj sati u semestru)	P	S	V	T
			15		45	
Status predmeta	Obavezni	Postotak primjene e-učenja	/			
OPIS PREDMETA						
Ciljevi predmeta	<p>Bioinformatika obuhvaća široko interdisciplinarno područje nastalo uslijed revolucionarnog napretka u razvoju molekularne biologije i računarstva početkom 21. stoljeća i potrebe da se da smisao i značenje te protumači velike količine podataka koje stvaraju nove visoko protočne metode analize bioloških makromolekula, kao što je područje genomike, proteomike, transkriptomike... Cilj bioinformatike je korištenjem baza podataka i računalnih algoritama analizirati pojedinačne gene, ili čitav DNA sadržaj nekog organizma, odnosno genom, pojedinačne proteine ili čitave kompleksne funkcionalne skupine makromolekula i biokemijskih putova. Kolegij je osmišljen s ciljem da se studenti upoznaju s nekima od osnovnih bioinformatičkih alata i bazama podataka povezanim s područjem transkriptomike i analize proteina kako bi u daljnjem radu mogli samostalno pristupiti osmišljavanju sličnih pokusa, prikupljanju podataka i njihovoj analizi, kao i samostalno nastaviti nadograđivanje znanja iz ovog iznimno dinamičnog i širokog područja.</p>					
Uvjeti za upis predmeta i ulazne kompetencije potrebne za predmet	Poznavanje osnova molekularne biologije i genetike.					
Očekivani ishodi učenja na razini predmeta (4-10 ishoda učenja)	<p>Po završetku kolegija, studenti bi trebali jasno znati:</p> <ul style="list-style-type: none"> <li>- Navesti glavne biološke baze podataka i kako ih pretraživati</li> <li>- Definirati BLAST algoritam, izvršiti uspoređivanje sekvenci s bazama podataka korištenjem istog algoritma i interpretirati rezultate</li> <li>- Definirati transkriptom i kako se proučava (RNAseq)</li> <li>- Ispravno osmisliti pokus kojem je cilj analiza transkriptoma</li> <li>- Koristiti neke od osnovnih bioinformatičkih alata za rukovanje podacima (UNIX naredbe, R paketi, RStudio)</li> <li>- Znati analizirati RNAseq pokus i protumačiti rezultate testiranja diferencijalne ekspresije gena</li> <li>- Navesti osnovne elemente građe i strukture proteina</li> <li>- Napraviti višestruko poravnavanje sekvenci proteina i dizajnirati proteinski model na temelju strukture homologa</li> <li>- Vizualizirati i analizirati protein bioinformatičkim alatima poput PyMOL-a, VMD-a i/ili sličnih programa koji se koriste za vizualizaciju makromolekula</li> </ul>					
Sadržaj predmeta detaljno razrađen prema satnici nastave	<p><b>Predavanja:</b></p> <p>1. Uvod u bioinformatiku i biološke baze podataka. (2 sata) Definicija bioinformatike i područja –omika istraživanja: genomika, transkriptomika, proteomika.... Pregled osnovnih baza na NCBI (National center for Biotechnology Information) portalu,</p>					

- EBI (European Bioinformatics Institute) i pristup sekvencama.
2. Pohrana sekvenci u FASTA formatu. BLAST algoritam (2 sata) Sanger sekvenciranje i definicija FASTA formata za pohranu sekvenci. Definicija i opis BLAST algoritma.
  3. Definicija, opis i regulacija transkriptoma (2 sata)  
Biologija transkripcije. Introni, egzoni, alternativno prekrajanje i izoforme. Sastav transkriptoma: kodirajuća i nekodirajuća RNA i njihova međusobna interakcija koja dinamički oblikuje transkriptom u svakom trenutku.
  4. Metode proučavanja transkriptoma sekvenciranjem - RNAseq (2 sata)  
Klasične metode proučavanja ekspresije gena i sastava transkriptoma u odnosu na sekvenciranje: prednosti i mane. Različiti tipovi knjižnica (*stranded vs unstranded, 3' tagged RNA-seq*). Različite vrste sekvenciranja: kratki i dugi sljedovi, upareni i neupareni sljedovi.
  5. Kako osmisлити dobar pokus za profiliranje transkriptoma (1 sat)  
Čimbenici na koje je potrebno misliti prilikom osmišljavanja pokusa, pripreme uzoraka za analizu i samog sekvenciranja kako bi se izbjegli nepoželjni serijski učinci koji mogu imati negativan učinak na analizu podataka. Prikaz svih koraka od postavljanja pokusa do analize podataka.
  6. Biologija translacije i struktura proteina. (2 sata)
  7. Jednostavno i multiplo poravnavanje. Homologni proteini (2 sata)
  8. Proteinske baze i modeliranje makromolekula (2 sata)

**Vježbe:**

1. Pretraživanje baza podataka. FASTA format i sređivanje sekvenci u programu MEGA. Usporedba sekvenci s bazama podataka korištenjem BLAST algoritma. Višestruki BLAST. (3 sata)
2. Uvod u UNIX (6 sati)  
Upoznavanje s radom u UNIX operativnom sustavu. Navigiranje kroz sustav (naredbe pwd, ls, cd), stvaranje, kopiranje i brisanje direktorija i dokumenata (mkdir, cp, rm), pregledavanje dokumenata (head, tail, cat, less), pretraživanje dokumenata (find), pretraživanje uzoraka (grep). Pokretanje procesa.
3. Upoznavanje s fastq formatom u kojem dolaze rezultati sekvenciranja (3 sata).  
Manipulirajte podacima u fastq formatu korištenjem unix naredbi.
4. Analiza kvalitete dobivenih sekvenci i podrezivanje sljedova u svrhu uklanjanja loših podataka (3 sata)  
Korištenjem programa FastQC studenti će provjeriti kvalitetu rezultata sekvenciranja i naučiti razlikovati dobre od loših podataka. Program Trimmomatic će biti korišten za uklanjanje adapterskih sekvenci i sljedova koji ne zadovoljavaju minimalne kriterije kvalitete.
5. Mapiranje sljedova na genom korištenjem aktualnih programa. Prikaz rezultata u SAM/BAM formatu i vizualizacija putem IGV-a (Integrative Genomics Viewer). Brojanje sljedova mapiranih iznad gena (3 sata)
6. Uvod u R (6 sati)  
Upoznavanje s programom R i RStudio za obradu podataka. Organizacija datoteka i stvaranje specifičnih projekata. Osnovni tipovi objekata u R-u (vektor, matrica, tablica, popis ili lista). Priprema podataka i uvođenje u R. Osnovni paketi za manipulaciju podacima (dplyr). Paketi okružja Bioconductor specifični za obradu genomskih podataka.

	<p>7. Diferencijalna ekspresija gena (6 sati) Eksplorativna analiza podataka u R-u (PCA, analiza glavnih komponenti). Normalizacija RNAseq podataka prema veličini knjižnice i veličini gena. Statistički modeli koji se koriste za modeliranje ekspresije gena temeljem negativne binomne raspodjele (R paketi EdgeR, DESeq2). Korekcije za višestruko testiranje i vizualizacija rezultata testa diferencijalne ekspresija gena u pokusu.</p> <p>8. Testiranje obogaćenja specifičnih genskih setova u skupu diferencijalno eksprimiranih gena: KEGG (Kyoto Encyclopedia of Genes and Genomes) i GO (Gene Ontology) genski setovi. Vizualizacija rezultata putem grafova i tablica (3 sata)</p> <p>9. Kako obraditi RNAseq podatke ako se radi o organizmu za kojega referentni genom još nije sastavljen? (2 sata) Uvod u <i>de novo assembly</i> transkriptoma (program Trinity), pronalaženje okvira čitanja i kodirajućih regija. Anotacija putem BLAST algoritma i osnovni programi za mapiranje sljedova na transkriptom (RSEM, Salmon).</p> <p>10. Jednostavno i multiplo poravnavanje. Dizajniranje homologa proteina pomoću višestrukog poravnanja sekvenci kao i identifikacija sekvence kandidata. Koristeći ExpASy alate izraditi homologni model. Translatirati model u sučelje za molekularnu vizualizaciju. Analiziranje strukture homolognog modela i identifikacija aminokiselinskih ostataka na važnim položajima u sekvenci proteina. (3 sata)</p> <p>11. Proteinske baze. ExpASy i PDB. Pretraživanje zadanih proteina i prikupljanje svih potrebnih podataka. (3 sata)</p> <p>12. Modeliranje makromolekula. Korištenjem ExpASy alata irada makromolekulskih modela i analiza njihove strukture. Identifikacija važnih aminokiselinskih ostataka i veza koje stvaraju unutar strukture. Lokalizacija aktivnog mjesta i ostataka važnih za katalizu. (4 sata)</p>					
Vrste izvođenja nastave:	<input checked="" type="checkbox"/> predavanja <input type="checkbox"/> seminari i radionice <input checked="" type="checkbox"/> vježbe <input type="checkbox"/> <i>on line</i> u cijelosti <input type="checkbox"/> mješovito e-učenje <input type="checkbox"/> terenska nastava		<input checked="" type="checkbox"/> samostalni zadaci <input checked="" type="checkbox"/> multimedija <input type="checkbox"/> laboratorij <input checked="" type="checkbox"/> mentorski rad <input type="checkbox"/> (ostalo upisati)			
Obveze studenata	<p>Nastava kolegija se organizira u obliku predavanja i vježbi (PowerPoint prezentacije, edukacijski filmovi, praktičan rad na računalima). Očekuje se da će studenti uredno pohađati nastavu i prisutnost na nastavi će se bilježiti. Od studenata se očekuje aktivno sudjelovanje u nastavnom procesu, izrada završnog projekta te polaganje usmenog ispita. U slučaju pojave okolnosti koje bi studente mogle spriječiti u izvršavanju svojih obaveza, očekuje se da će na vrijeme o tome obavijestiti predmetne nastavnike.</p>					
Praćenje rada studenata ( <i>upisati udio u ECTS bodovima za svaku aktivnost tako da ukupni broj ECTS bodova odgovara bodovnoj vrijednosti predmeta</i> ):	Pohađanje nastave	1	Istraživanje		Praktični rad	2
	Eksperimentalni rad		Referat		(Ostalo upisati)	
	Esej		Seminarski rad		(Ostalo upisati)	
	Kolokviji		Usmeni ispit	1	(Ostalo upisati)	
	Pismeni ispit		Projekt	1	(Ostalo upisati)	

<p>Ocjenjivanje i vrjednovanje rada studenata tijekom nastave i na završnom ispitu</p>	<p>Prisutnost studenata tijekom nastave će se evidentirati svaki sat. Studenti će se poticati na aktivno sudjelovanje u nastavi i samostalno rješavanje problema i zadataka što će se periodički bodovati. Na kraju kolegija studenti će odabrati projekt kroz koji će analizirati i riješiti problem iz područja transkriptomike ili analize proteina koji ih zanima. Studenti će samostalno i savjetujući se s nastavnicima izraditi projekt i izložiti ga nastavnicima na ispitnom roku u sklopu usmenog ispita. Ukupna ocjena će uključivati ocjenu ukupnog zalaganja studenta tijekom kolegija, aktivnosti na nastavi, provedbe projekta i diskusije na usmenom ispitu.</p> <p>Vrednovanje i konačna ocjena:  90% - 100%: ocjena 5 (izvrstan);  80% - 89%: ocjena 4 (vrlo dobar);  65% - 79%: ocjena 3 (dobar);  51% - 64%: ocjena 2 (dovoljan);  &lt; 51%: ocjena 1 (nedovoljan).</p>		
<p>Obvezna literatura (dostupna u knjižnici i putem ostalih medija)</p>	<p><b>Naslov</b></p>	<p><b>Broj primjeraka u knjižnici</b></p>	<p><b>Dostupnost putem ostalih medija</b></p>
	<p>Bioinformatics and Functional Genomics, third edition. 2015. Jonathan Pevsner. Wiley-Blackwell.</p>	<p>1</p>	
	<p>Introduction to Bioinformatics, 3rd Edition. 2008. Arthur M. Lesk. Oxford University Press.</p>	<p>1</p>	
<p>Dopunska literatura</p>	<p>1. Love MI, Anders S, Kim V and Huber W. RNA-Seq workflow: gene-level exploratory analysis and differential expression [version 2; peer review: 2 approved]. F1000Research 2016, 4:1070 (<a href="https://doi.org/10.12688/f1000research.7035.2">https://doi.org/10.12688/f1000research.7035.2</a>)  2. Zuur AF, Ieno EN, Meesters EHWG. . A beginner's guide to R. 2009. Springer, New York.  3. Trumbić Ž. 2019. Interna skripta – vježbe iz kolegija Bioinformatika. Sveučilište u Splitu.  4. Trumbić Ž. 2019. Dodatni nastavni materijal - ppt predavanja i popratni materijal za vježbe iz kolegija Bioinformatika.</p>		
<p>Načini praćenja kvalitete koji osiguravaju stjecanje utvrđenih ishoda učenja</p>	<p>Predavanja i vježbe su zamišljene kao interaktivna aktivnost te se potiče aktivno sudjelovanje studenata u nastavi kroz pitanja, komentare i samostalno rješavanje zadataka čime se omogućava kontinuirano praćenje rada studenata. Analizom uspjeha kroz završni projekt i usmeni ispit će se pratiti razina stjecanja utvrđenih ishoda učenja.</p>		
<p>Ostalo (prema mišljenju predlagatelja)</p>	<p>Vrijeme konzultacija i termini ispita će biti oglaseni na internet stranicama Prirodoslovno-matematičkog fakulteta Split. Studenti mogu kontaktirati nastavnike putem e-maila na <a href="mailto:msprung@pmfst.hr">msprung@pmfst.hr</a> i <a href="mailto:ztrumbic@unist.hr">ztrumbic@unist.hr</a> za sva pitanja i pomoć pri savladavanju sadržaja kolegija.</p>		