

KARTA KURSU

Nazwa	Bioinformatyka sekwencji biologicznych
Nazwa w j. ang.	Sequence bioinformatics

Koordynator	Dr Grzegorz Migdałek	Zespół dydaktyczny
Punktacja ECTS*	4	

Opis kursu (cele kształcenia)

Celem kursu jest prezentacja podstaw teoretycznych oraz metod obliczeniowych wykorzystywanych do analiz sekwencji biologicznych. Program kształcenia obejmuje:

1. Przedstawienie strategii składowania i zarządzania danymi sekwencyjnymi przy pomocy narzędzi i technologii informatycznych.
2. Przekazanie wiedzy na temat dostępnych zasobów oraz rodzajów danych biologicznych.
3. Prezentacja teoretyczna i praktyczna technik oraz strategii przeszukiwania biologicznych baz danych.
4. Wykształcenie umiejętności analizy i efektywnego pozyskiwania informacji zawartych w rekordach danych biologicznych.
5. Przedstawienie podstaw działania algorytmów obliczeniowych wykorzystywanych do porównywania sekwencji.
6. Wykształcenie umiejętności prawidłowego i wydajnego wykorzystywania narzędzi bioinformatycznych do porównywania sekwencji biologicznych.
7. Prezentacja wybranych technik identyfikacji genów i ich opisu funkcjonalnego.
8. Przekazanie wiedzy na temat metod i algorytmów stosowanych do identyfikacji motywów w sekwencjach biologicznych.
9. Przedstawienie metod wyszukiwania sekwencji spokrewnionych ewolucyjnie oraz wykształcenie umiejętności interpretacji otrzymywanych wyników.

Warunki wstępne

Wiedza	Znajomość podstawowych pojęć i procesów z zakresu genetyki, programowania i bioinformatyki
Umiejętności	Podstawowe umiejętności posługiwania się systemem operacyjnym Linux
Kursy	Podstawy oprogramowania, Wstęp do programowania, Wprowadzenie do bioinformatyki, Algorytmy i struktury danych, Programy do analizy danych

Efekty uczenia się

	Efekt uczenia się dla kursu	Odniesienie do efektów kierunkowych
Wiedza	W01, ma wiedzę na temat dostępnych zasobów oraz rodzajów danych biologicznych	K_W01, K_W11
	W02, posiada wiedzę na temat strategii składowania i zarządzania danymi sekwencyjnymi oraz przeszukiwania biologicznych baz danych.	K_W11, K_W17
	W03, zna podstawy działania algorytmów obliczeniowych wykorzystywanych do porównywania sekwencji.	K_W09, K_W10, K_W17
	W04, ma wiedzę na temat wybranych technik identyfikacji genów i ich opisu funkcjonalnego.	K_W03, K_W05, K_W17

	Efekt uczenia się dla kursu	Odniesienie do efektów kierunkowych
Umiejętności	U01, analizuje i efektywnie pozyskuje informacje zawarte w rekordach danych biologicznych.	K_U01, K_U02
	U02, prawidłowo i wydajnie wykorzystuje narzędzia bioinformatyczne do porównywania sekwencji biologicznych.	K_U05, K_U06, K_U09
	U03, wykorzystuje techniki wyszukiwania sekwencji spokrewnionych ewolucyjnie oraz interpretuje otrzymywane wyniki.	K_U02, K_U06

	Efekt uczenia się dla kursu	Odniesienie do efektów kierunkowych
Kompetencje społeczne	K01. Rozumie konieczność uzupełniania wiedzy w oparciu o aktualne dane	K_K01
	K02. W interpretacji zjawisk korzysta z podstaw empirycznych oraz metod statystycznych i narzędzi informatycznych	K_K03

Organizacja												
Forma zajęć	Wykład (W)	Ćwiczenia w grupach										
		A		K		L		S		P		E
Liczba godzin	10					30						

Opis metod prowadzenia zajęć

Podczas wykładów omawiane są realizowane treści, przedstawiane są schematy, zdjęcia, rysunki oraz prezentacje multimedialne.

Podczas ćwiczeń studenci zapoznają się w praktyce z metodami przeszukiwania baz danych sekwencji nukleotydowych i białkowych oraz algorytmami dopasowywania dwóch i wielu sekwencji z użyciem specjalistycznego oprogramowania i samodzielnie tworzonych skryptów

Formy sprawdzania efektów kształcenia

	E – learning	Gry dydaktyczne	Ćwiczenia w szkole	Zajęcia terenowe	Praca laboratoryjna	Projekt indywidualny	Projekt grupowy	Udział w dyskusji	Referat	Praca pisemna (esej)	Egzamin ustny	Egzamin pisemny	Inne
W01						X						X	
W02						X						X	
W03						X						X	
W04						X						X	
U01					X	X							
U02					X	X							
U03					X	X							
K01					X	X							
K02					X	X							

Kryteria oceny	Warunkiem pozytywnego zaliczenia kursu jest zaliczenie ćwiczeń (końcowy projekt indywidualny) oraz pozytywna ocena z egzaminu pisemnego (50% +1 punktów)
----------------	--

Uwagi	Kurs prowadzony w języku polskim
-------	----------------------------------

Treści merytoryczne (wykaz tematów)

1. Edycja surowych danych sekwencyjnych i podstawowe formaty danych sekwencji (ab1, scf, FASTA)
2. Formaty danych sekwencji w bazach danych (NCBI, EBI), narzędzia do wyszukiwania i pobierania danych.
3. Strategie wyszukiwania informacji w biologicznych bazach danych z wykorzystaniem systemów ENTREZ (NCBI) i SRS (EBI).
4. Rekordy sekwencji i operacje na sekwencjach w środowisku Python i R, statystyki pojedynczych sekwencji.
5. Metody obliczeniowe stosowane do lokalnego i globalnego dopasowywania dwóch sekwencji (matryce punktowe, programowanie dynamiczne, tablice podstawień).
6. Przegląd metod statystycznych stosowanych do oznaczania jakości dopasowania dwóch sekwencji.
7. Programy BLAST i FASTA oraz praktyczne aspekty przeszukiwania baz danych sekwencji biologicznych w oparciu o podobieństwo.
8. Algorytmy progresywne i iteracyjne dopasowania wielu sekwencji (ClustalW, MUSCLE, MAFFT, T-coffee), wyszukiwanie bloków konserwatywnych (Gblocks)
9. Algorytmy dopasowywania długich sekwencji (MUM, LCS, MSA) i fragmentów sekwencji (tablice haszujące, transformanta Burrowsa-Wheelera)
10. Porównywanie sekwencji bez dopasowania (frekwencje k-merów, metody złożoności i entropii)
11. Podstawy adnotacji funkcjonalnej sekwencji biologicznych (wyszukiwanie motywów sekwencji; metody przewidywania sekwencji kodujących białka; wyszukiwanie promotorowych motywów regulatorowych; identyfikacja domen białkowych)

Wykaz literatury podstawowej

Higgs P.G., Attwood T.K. (2008) Bioinformatyka i ewolucja molekularna. Wydawnictwo Naukowe PWN.
 Polański A. (2010) Podstawy Bioinformatyki. Wydawnictwo PJWSTK.

Wykaz literatury uzupełniającej

Baxevanis A.D., Ouellette B.F. (2005) Bioinformatyka. Podręcznik do analizy genów i białek. Wydawnictwo Naukowe PWN.
 Mount D.W. (2004) Bioinformatics: sequence and genome analysis. Cold Spring Harbor Laboratory Press

Bilans godzinowy zgodny z CNPS (Całkowity Nakład Pracy Studenta)

liczba godzin w kontakcie z prowadzącymi	Wykład	10
	Konwersatorium (ćwiczenia, laboratorium itd.)	30
	Pozostałe godziny kontaktu studenta z prowadzącym	5
liczba godzin pracy studenta bez kontaktu z prowadzącymi	Lektura w ramach przygotowania do zajęć	5
	Przygotowanie krótkiej pracy pisemnej lub referatu po zapoznaniu się z niezbędną literaturą przedmiotu	0
	Przygotowanie projektu lub prezentacji na podany temat (praca w grupie)	5
	Przygotowanie do egzaminu/zaliczenia	5
Ogółem bilans czasu pracy		60
Liczba punktów ECTS w zależności od przyjętego przelicznika		4