

--	--	--

1. Nazwa przedmiotu: Kompresja i analiza danych genomowych		2. Kod przedmiotu: KiADG		
3. Karta przedmiotu ważna od roku akademickiego: 2017/2018				
4. Forma kształcenia: studia doktoranckie				
5. Forma studiów: studia niestacjonarne/stacjonarne				
6. Kierunek studiów: INFORMATYKA				
7. Profil studiów: akademicki				
8. Specjalność:				
9. Semestr:				
10. Jednostka prowadząca przedmiot: Instytut Informatyki, RAu2				
11. Prowadzący przedmiot: dr inż. Agnieszka Danek				
12. Przynależność do grupy przedmiotów: przedmioty fakultatywne				
13. Status przedmiotu:				
14. Język prowadzenia zajęć: polski/angielski				
15. Przedmioty wprowadzające oraz wymagania wstępne: Bioinformatyka. Znajomość podstawowych zagadnień oraz terminologii z zakresu bioinformatyki.				
16. Cel przedmiotu: Celem przedmiotu jest zrozumienie metodologii tworzenia narzędzi oraz serwisów porównawczej analizy sekwencji biologicznych jak również poznanie ich zalet oraz ograniczeń. Studenci zostaną przygotowani do samodzielnego tworzenia procedur analizy danych oraz wykorzystania narzędzi bioinformatycznych do rozwiązywania problemów badawczych z dziedziny genomiki.				
17. Efekty kształcenia: ¹				
Nr	Opis efektu kształcenia	Metoda sprawdzenia efektu kształcenia	Forma prowadzenia zajęć	Odniesienie do efektów dla kierunku studiów
W1	Ma wiedzę dotyczącą najważniejszych problemów bioinformatycznych oraz metodologii ich rozwiązywania		WT, WM	RAU_I_W01
W2	Zna zaawansowane narzędzia i serwisy analizy porównawczej sekwencji biologicznych		WT, WM	RAU_I_W03
W3	Ma wiedzę dotyczącą technologii zarządzania danymi biologicznymi oraz oprogramowaniem bioinformatycznym		WT, WM	RAU_I_W05
W4	Zna techniki projektowania narzędzi bioinformatycznych		WT, WM	RAU_I_W07
U1	Potrafi sformułować problemy badawcze oraz tworzyć zaawansowane procedury analizy danych		WT, WM	RAU_I_U01
U2	Posiada umiejętność wyszukiwania, przetwarzania oraz selekcjonowania informacji z różnych źródeł		WT, WM	RAU_I_U03
U3	Posiada umiejętność analizy oraz interpretacji otrzymywanych wyników badań		WT, WM	RAU_I_U06
U4	Posiada umiejętność projektowania algorytmów służących rozwiązaniu określonych problemów badawczych w genomice		WT, WM	RAU_I_U04
K1	Potrafi pracować i komunikować się w zespole		WT, WM	RAU_I_K01
K2	Rozumie potrzebę aktualizacji wiedzy dotyczącej metod i narzędzi badawczych oraz wykorzystywania jej w działalności naukowej		WT, WM	RAU_I_K02

¹ należy wskazać ok. 5 – 8 efektów kształcenia

18. Formy zajęć dydaktycznych i ich wymiar (liczba godzin) W. : 10 L.: 0
19. Treści kształcenia: Wykład Wprowadzenie do przedmiotu: podstawowe pojęcia. Metody oraz narzędzia modyfikacji sekwencji genomowych. Wyszukiwanie otwartych ramek odczytu, ORF. Metody badania podobieństwa oraz wzajemnego dopasowania sekwencji biologicznych. Algorytmy dopasowania dwóch sekwencji. Przeszukiwanie baz danych sekwencji biologicznych. Algorytmy BLAST oraz FASTA. Ocena jakości dopasowania. Metody dopasowania wielu sekwencji biologicznych. Zasada działania programu Clustal. Edycja wzajemnego dopasowania wielu sekwencji DNA. Analiza filogenetyczna. Podstawowe pojęcia, topologia drzew filogenetycznych. Metody konstruowania molekularnych drzew filogenetycznych, algorytm średnich połączeń UPGMA, metoda najbliższego sąsiedztwa, metoda bootstrap. Narzędzia do asemblacji sekwencji DNA. Metody składania pełnych sekwencji genomowych z krótkich odczytów. Mapowanie krótkich odczytów sekwencji do genomu referencyjnego. Struktura drugorzędowa RNA. Podstawowe pojęcia, formy prezentacji struktury drugorzędowej. Metody przewidywania struktury drugorzędowej RNA. Algorytm Nussinova optymalizacji ilości par zasad. Algorytm Zukera optymalizacji energii swobodnej struktury RNA. Wyszukiwanie charakterystycznych motywów w sekwencjach biologicznych. Metody opisu i prezentacji motywów. Zasada działania programów do wyszukiwania charakterystycznych motywów w sekwencjach DNA. Projekt 1000 genomów. Założenia oraz obecny status projektu. Formaty przechowywania danych. Metody dostępu do danych z projektu 1000 genomów. Narzędzia analizy oraz przetwarzania danych z projektu. Zajęcia praktyczne Uzupełnieniem wykładów z przedmiotu są zajęcia praktyczne, w ramach których studenci mają możliwość praktycznego zapoznania się z omawianymi metodami oraz narzędziami stosowanymi w genomice. 1. Przeszukiwanie biologicznych baz danych. System Entrez – metody dostępu oraz przeszukiwania najważniejszych baz danych sekwencji molekularnych. Zasada działania algorytmów BLAST oraz FASTA. Ocena jakości dopasowania sekwencji. 2. Wzajemne dopasowanie wielu sekwencji. Zasada działania programu Clustal. 3. Analiza filogenetyczna. Metody konstruowania molekularnych drzew filogenetycznych. 4. Składanie sekwencji genomowych z krótkich odczytów. Mapowanie krótkich odczytów sekwencji do genomu referencyjnego. 5. Przewidywanie struktury drugorzędowej RNA. Algorytm Zukera optymalizacji energii swobodnej struktury RNA. Zasada działania serwisu RNAfold. 6. Wyszukiwanie charakterystycznych motywów w sekwencjach biologicznych. Metody opisu i prezentacji motywów. Zasada działania programu MEME. 7. Projekt 1000 genomów. Dostęp do danych z projektu 1000 genomów. Narzędzia przetwarzania oraz analizy danych z projektu.
20. Egzamin: nie Zaliczenie na podstawie dyskusji w trakcie wykładów oraz opcjonalnie na podstawie przedłożonego przez studenta opracowania pisemnego wybranego zagadnienia.
21. Literatura podstawowa: 1. A. D. Baxevanis, B. F. F. Ouellette. Bioinformatyka: podręcznik do analizy genów i białek. PWN, 2005. 2. Paul G. Higgs, Teresa K. Attwood. Bioinformatyka i ewolucja molekularna. PWN, 2008. 3. R. C. Deonier, S. Tavaré, and M. S. Waterman. Computational Genome Analysis: An Introduction. Springer, 2005. 4. C. Gibas, and P. Jambeck. Developing Bioinformatics Computer Skills. O'Reilly Media, 2001.
22. Literatura uzupełniająca: 1. H.-J. Bockenhauer, D. Bongartz. Algorithmic Aspects of Bioinformatics. Springer, 2007. 2. D. W. Mount. Bioinformatics: Sequence and Genome Analysis. Cold Spring Harbor Press, 2001. 3. A. Polanski, and M. Kimmel. Bioinformatics. Springer, 2007. 4. D. Gusfield. Algorithms on Strings, Trees, and Sequences. Cambridge University Press, 1997.

23. Nakład pracy studenta potrzebny do osiągnięcia efektów kształcenia

Lp.	Forma zajęć	Liczba godzin kontaktowych / pracy studenta
1	Wykład	10/5
2	Ćwiczenia	0/0
3	Laboratorium	0/0
4	Projekt	0/0
5	Seminarium	0/0
6	Inne: konsultacje	5/5
	Suma godzin	15/10

24. Suma wszystkich godzin: 25**25. Liczba punktów ECTS:² 1****26. Liczba punktów ECTS uzyskanych na zajęciach z bezpośrednim udziałem nauczyciela akademickiego: 1****27. Liczba punktów ECTS uzyskanych na zajęciach o charakterze praktycznym (laboratoria, projekty): 0**

26. Uwagi: Efekty kształcenia w zakresie wiedzy weryfikowane są na bieżąco w trakcie wykładów, natomiast umiejętności podlegają weryfikacji poprzez formułowanie i rozwiązywanie zadań praktycznych. Efekty kształcenia w zakresie kompetencji społecznych sprawdzane są w trakcie pracy zespołowej nad przykładowymi problemami badawczymi oraz przy opracowywaniu i prezentacji raportów końcowych.

Zatwierdzono:

.....
(data i podpis kierownika studiów doktoranckich)

² 1 punkt ECTS – 25-30 godzin.