

## Computationele biologie (C003789)

**Wegens Covid19 kan mogelijk afgeweken worden van de onderwijs- en evaluatievormen. Dergelijke afwijkingen zullen via Ufora worden gecommuniceerd.**

**Cursusomvang** *(nominale waarden; effectieve waarden kunnen verschillen per opleiding)*

**Studiepunten 6.0**      **Studietijd 180 u**      **Contacturen**      60.0 u

**Aanbodsessies en werkvormen in academiejaar 2022-2023**

A (semester 2)	Nederlands	Gent	werkcollege: PC- klasoefeningen	20.0 u
			hoorcollege	30.0 u
			project	12.5 u

**Lesgevers in academiejaar 2022-2023**

Dawyndt, Peter      WE02      Verantwoordelijk lesgever

**Aangeboden in onderstaande opleidingen in 2022-2023**

<a href="#">Bachelor of Science in de informatica</a>	<b>stptn</b>	<b>aanbodssessie</b>
	6	A

**Onderwijstalen**

Nederlands

**Trefwoorden**

Rekenmethoden, moleculaire biologie, genoomstructuur, genpredictie, sequenties aligneren, fylogenie, vergelijkend genoomonderzoek, analyse van genexpressie

**Situering**

Waar komt SARS vandaan? Hebben wij genen geërfd van de Neanderthalers? Hoe gebruiken planten hun interne klok? Hoeveel genetisch materiaal wordt er uitgewisseld tussen bacteriële soorten? De genomische revolutie in de biologie maakt het mogelijk om een antwoord te vinden op elk van deze vragen. Maar deze revolutie zou onmogelijk geweest zijn zonder de ondersteuning van krachtige computationele en statistische methodes die toelaten om gebruik te maken van deze genomische informatie (1.3). Oplossen van openstaande en toekomstige problemen die opduiken in dit intrigerend onderzoeksdomein (1.2) vereist de opleiding van een volgende generatie onderzoekers, die zich vloeiend kunnen uitdrukken in de taal van de wiskunde, informatica en biologie (3.1,3.2,3.5).

Ontdek de wereld van de computationele biologie doorheen een aantal case studies die zijn uitgewerkt in Matlab/Python. Deze case studies stellen de student in staat om zichzelf vertrouwd te maken met de terminologie en de basisprincipes van de moleculaire biologie, en bieden een overzicht van de waaier aan computationele technieken die gebruikt worden in dit groeiende onderzoeksdomein (2.1,2.2,2.4). Dit opleidingsonderdeel effent de weg om vlot dit onderzoeksgedebied te kunnen binnenstappen (1.6,5.1,5.5). Het loodst de studenten langs te belangrijkste verwezenlijkingen van de bioinformatica, gebaseerd op een praktijkgerichte benadering. Statistische sequentie-analyse, alignering van sequenties, hidden Markov modellen, genpredictie en andere zaken (1.1) worden op een nauwgezette maar toegankelijke manier aangebracht.

**Inhoud**

- Inleiding tot de celbiologie
- Statistische eigenschappen van sequenties
- Genpredictie

- Sequenties aligneren
- Probabilistische sequentiemodellen (hidden Markov models)
- Genetische variatie binnen en tussen soorten
- Evolutie: natuurlijke selectie op moleculair niveau
- Fylogenetische analyse
- Vergelijking van complete genomen
- Genexpressiepatronen
- Identificatie van regulerende sequenties

### Begincompetenties

Basiskennis van waarschijnlijkheidstheorie, algoritmische concepten en programmeervaardigheden zijn essentieel. Beheersen van de programmeertaal Python en kunnen werken met wetenschappelijke modules (numpy, scipy, matplotlib) zijn vereisten. Er wordt geen voorkennis van de moleculaire biologie verondersteld, aangezien doorheen dit opleidingsonderdeel voldoende tijd zal worden uitgetrokken om de biologische achtergrond te schetsen. Het is noodzakelijk om over een gezonde dosis interesse in de biologische aspecten van het probleemdomein te beschikken.

### Eindcompetenties

- 1 Inzicht in statistische en algoritmische strategieën voor sequentie- en genexpressie-analyse.
- 2 Aanvoelen hoe en waarom wiskunde en informatica een rol kan spelen in het hedendaags biologisch onderzoek.
- 3 Kunnen lezen en begrijpen van recente wetenschappelijke artikels over genomische onderwerpen, en in staat zijn om de methodes voor gegevensverwerking die erin voorkomen te beoordelen.
- 4 Vertrouwd zijn met de standaard technieken en problemen, én de huidige en toekomstige uitdagingen uit het onderzoeksdomein van de computationele biologie.
- 5 Verwerven en verwerken van genomische onderzoeksgegevens.

### Creditcontractvoorwaarde

Toelating tot dit opleidingsonderdeel via creditcontract is mogelijk mits gunstige beoordeling van de competenties

### Examencontractvoorwaarde

Dit opleidingsonderdeel kan niet via examencontract gevolgd worden

### Didactische werkvormen

Hoorcollege, project, werkcollege: PC-klasoefeningen

### Toelichtingen bij de didactische werkvormen

Begeleide werkcolleges in de PC-zaal, waarbij de uitgevoerde analyses periodiek gerapporteerd moeten worden.

### Leermateriaal

Handboek: Compeau P, Pevzner P (2015). Bioinformatics algorithms: an active learning approach. Active Learning Publishers. 2 volumes: ISBN-13:978-0990374619 and ISBN-13:978-0990374626.

Rekencluster waarop de nodige software voorgeïnstalleerd is

Bijkomend leermateriaal wordt ter beschikking gesteld via de elektronische leeromgeving Ufora

- diavoorstelling in Powerpoint uit de hoorcolleges
- (publiek) beschikbare software tools
- wetenschappelijke tijdschriften met betrekking tot computationele biologie
- informatieve websites

Geraamde totaalprijs: 0 EUR

### Referenties

Durbin R., S. Eddy, A. Krogh en G. Mitchinson. Biological Sequence Analysis - Probabilistic models of proteins and nucleic acids, Cambridge University Press, 1998. (ISBN-13: 978-0521629713 | ISBN-10: 0-521-62971-3)

Felsenstein J., Inferring Phylogenies, Sinauer Associates, 2003. (ISBN-13: 978-0878931774 | ISBN-10: 0-878-93177-5)

Gusfield D., Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology, Cambridge University Press, 1997. (ISBN-13: 978-0521585194 | ISBN-10: 0-521-58519-8)

Jones N.C. en P.A. Pevzner, An Introduction to Bioinformatics Algorithms, MIT Press, Cambridge, MA, 2004. (ISBN-13: 978-0262101066 | ISBN-10: 0-262-10106-8)

Setubal C. en J. Meidanis, Introduction to Computational Molecular Biology, PWS Publishing,

(Goedgekeurd)

1997. (ISBN-13: 978-0534952624 | ISBN-10: 0-534-95262-3)

Mount D.W., Bioinformatics: Sequence and Genome Analysis. Cold Spring Harbor Laboratory Press, 2001. (ISBN-13: 978-0879697129 | ISBN-10: 0-879-69712-1)

### **Vakinhoudelijke studiebegeleiding**

Algoritmische aspecten van de computationele biologie die tijdens de hoorcolleges aan bod komen, worden geïllustreerd aan de hand van een reeks case studies die uitgewerkt zijn in Matlab of Python. Sommige case studies komen tijdens de hoorcolleges aan bod, terwijl andere case studies onder begeleiding zelfstudie door de student moeten worden uitgewerkt tijdens werkcolleges (in computerklas) of onder de vorm van verplichte projecten. Centraal hierbij staat een doorgedreven gebruik van de elektronische leeromgeving Ufora, waarbij feedback en bijkomend leermateriaal kan worden uitgewisseld tussen de begeleiders en de studenten onderling.

### **Evaluatiemomenten**

niet-periodegebonden evaluatie

### **Evaluatievormen bij periodegebonden evaluatie in de eerste examenperiode**

### **Evaluatievormen bij periodegebonden evaluatie in de tweede examenperiode**

### **Evaluatievormen bij niet-periodegebonden evaluatie**

Portfolio, werkstuk, verslag

### **Tweede examenkans in geval van niet-periodegebonden evaluatie**

Examen in de tweede examenperiode is mogelijk

### **Toelichtingen bij de evaluatievormen**

Studenten worden geëvalueerd door een reeks van programmeeroefeningen die als verplicht huiswerk moeten uitgewerkt worden (niet-periodegebonden evaluatie: 60%) en een individueel project (periodegebonden evaluatie: 40%).

### **Eindscoreberekening**

Niet-periode gebonden evaluatie (60%) en periodegebonden project-evaluatie (40%). Minstens de helft van de punten moet behaald worden voor beide onderdelen: niet-periodegebonden evaluatie en project-evaluatie. Als slechts voor één van beide onderdelen minstens de helft van de punten behaald wordt, dan wordt de eindscore enkel gebaseerd op het andere onderdeel, dat dan gewicht 100% krijgt.

Voor de tweede zitting worden voor de vier opdrachten alternatieve opgaven aangeduid (andere dan diegene uit eerste zitting). Het project blijft gewoon hetzelfde. Studenten kunnen elk van deze opdrachten/project hernemen in tweede zitting. Hernemen ze een opdracht niet, dan behouden ze hun score voor die opdracht uit eerste zitting.