



ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN

**ĐỀ CƯƠNG MÔN HỌC**

**TIN - SINH HỌC**

**1. THÔNG TIN CHUNG (General information)**

Tên môn học (tiếng Việt):	Tin - Sinh học
Tên môn học (tiếng Anh):	BioInformatics
Mã môn học:	IE401
Thuộc khối kiến thức:	Đại cương <input type="checkbox"/> ; Cơ sở nhóm ngành <input type="checkbox"/> ; Cơ sở ngành <input type="checkbox"/> ; Chuyên ngành <input type="checkbox"/> ; Tốt nghiệp <input checked="" type="checkbox"/>
Khoa, Bộ môn phụ trách:	Khoa học và Kỹ thuật Thông tin
Giảng viên biên soạn:	ThS. Nguyễn Thành Trung Email: trungnt@uit.edu.vn
Số tín chỉ:	3
Lý thuyết:	3
Thực hành:	0
Tự học:	6
Môn học tiên quyết:	Không
Môn học trước:	Cấu trúc Dữ liệu và Giải thuật

**2. MÔ TẢ MÔN HỌC (Course description)**

Tin sinh học là một lĩnh vực khoa học sử dụng các công nghệ của các ngành: toán học ứng dụng, tin học, thống kê, khoa học máy tính, trí tuệ nhân tạo, hóa học và hóa sinh để giải quyết các vấn đề sinh học. Nội dung Tin sinh học bao gồm: Sinh học phân tử. Giới thiệu một số ngân hàng cơ sở dữ liệu sinh học trên Internet. Thuật toán giải một số bài toán cơ bản trong Tin sinh học. Xây dựng phần mềm Tin sinh học.

### 3. MỤC TIÊU MÔN HỌC (Course goals)

Sau khi hoàn thành môn học này, sinh viên có thể:

Bảng 1.

Mục tiêu môn học
Sinh viên nắm vững nền tảng kiến thức, cơ sở lý luận và phương pháp nghiên cứu cơ bản.
Sinh viên có khả năng khai thác và sử dụng các công cụ trực tuyến, các phần mềm hỗ trợ trong việc phân tích và xử lý các thông tin sinh học.
Hiểu được cây phát sinh loài. Biết so sánh các trình tự, phân tích trình tự.
Có khả năng phân tích các trình tự sinh học bằng các thuật toán tin học, mức độ biểu hiện gen, nhận diện protein, dự đoán cấu trúc protein.
Có khả năng truy cập đến các CSDL sinh học lớn để tìm kiếm.
Có khả năng xây dựng các phần mềm đơn giản giải quyết một số bài toán Tin sinh học.

### 4. CHUẨN ĐẦU RA MÔN HỌC (Course learning outcomes)

Bảng 2.

CĐRMH	Mô tả CĐRMH (Mục tiêu cụ thể)
G1 (LO 2, LO 3)	Sinh viên nắm vững nền tảng kiến thức, cơ sở lý luận và phương pháp nghiên cứu cơ bản.
G2 (LO 2, LO 3)	Sinh viên có khả năng khai thác và sử dụng các công cụ trực tuyến, các phần mềm hỗ trợ trong việc phân tích và xử lý các thông tin sinh học.
G3 (LO 2, LO 3, LO 5)	Hiểu được cây phát sinh loài. Biết so sánh các trình tự, phân tích trình tự.
G4 (LO 5, LO 6)	Có khả năng phân tích các trình tự sinh học bằng các thuật toán tin học, mức độ biểu hiện gen, nhận diện protein, dự đoán cấu trúc protein.
G5 (LO 5, LO 10)	Có khả năng truy cập đến các CSDL sinh học lớn để tìm kiếm.
G6 (LO 6, LO 7, LO 10)	Có khả năng xây dựng các phần mềm đơn giản giải quyết một số bài toán Tin sinh học.

### 5. NỘI DUNG MÔN HỌC, KẾ HOẠCH GIẢNG DẠY (Course content, lesson plan)

Lý thuyết

Bảng 3.

<b>Buổi học (45 tiết)</b>	<b>Nội dung</b>	<b>CĐRMH</b>	<b>Thành phần đánh giá</b>
Buổi 1 (4 tiết)	<b>Chương 1: Giới thiệu về Tin sinh học</b> <ol style="list-style-type: none"> <li>1. Khái niệm</li> <li>2. Sự ra đời của tin sinh học</li> <li>3. Vị trí của tin sinh học trong nghiên cứu và ứng dụng</li> <li>4. Nhiệm vụ của tin sinh học</li> </ol>	G1	A2, A4
Buổi 2-4 (9 tiết)	<b>Chương 2: Cơ sở của Sinh học tính toán</b> <ol style="list-style-type: none"> <li>1. Giới thiệu so sánh trình tự và lập trình tin sinh học</li> <li>2. Bắt cặp nhiều trình tự (Multiple Sequence Alignments)</li> <li>3. Phân tích hệ phát sinh loài (Phylogenetic)</li> <li>4. Giải trình tự Genome và phân tích trình tự DNA</li> <li>5. So sánh và bắt cặp trình tự DNA</li> <li>6. Tìm và mô phỏng DNA motif</li> <li>7. Mô hình Markov và Markov ẩn cho trình tự DNA</li> <li>8. Sự tiến hóa trình tự DNA</li> </ol>	G2, G3, G4	A2, A4
Buổi 5-7 (8 tiết)	<b>Chương 3: Tổng quan về sinh học phân tử</b> <ol style="list-style-type: none"> <li>1. Các đại phân tử sinh học và cấu trúc của chúng</li> <li>2. Luận thuyết trung tâm</li> <li>3. Biểu diễn dữ liệu sinh học phân tử trên máy tính</li> </ol>	G3, G4	A4
Buổi 8-10 (8 tiết)	<b>Chương 4: Giới thiệu CSDL sinh học và các công cụ phân tích</b> <ol style="list-style-type: none"> <li>1. Cơ sở dữ liệu sinh học (ngân hàng sinh học)</li> <li>2. Giới thiệu công cụ phân tích cơ sở dữ liệu sinh học trực tuyến</li> <li>3. Giới thiệu phần mềm Tin sinh học</li> </ol>	G3	A4
Buổi 11-13 (10 tiết)	<b>Chương 5: Ứng dụng một số công cụ trực tuyến trong việc phân tích trình tự sinh học</b>	G3, G4, G5, G7, G8	A4

	<ol style="list-style-type: none"> <li>1. Tìm kiếm các trình tự ADN, axit amin trong ngân hàng dữ liệu</li> <li>2. Tìm kiếm các trình tự ADN, Protein tương đồng</li> <li>3. So sánh trình tự (Sequence Alignment)</li> <li>4. Đăng ký trình tự sinh học vào các ngân hàng cơ sở dữ liệu sinh học</li> </ol>		
Buổi 14, 15 (6 tiết)	<b>Chương 6: Xây dựng phần mềm tin sinh học</b> <ol style="list-style-type: none"> <li>1. Giới thiệu một số bài toán cơ bản trong Tin sinh học</li> <li>2. Sử dụng ngôn ngữ lập trình Perl viết chương trình giải một số bài toán cơ bản trong Tin sinh học</li> </ol>	G3, G4, G5, G6, G7, G8	A4

## 6. ĐÁNH GIÁ MÔN HỌC (Course assessment)

Bảng 4.

Thành phần đánh giá	CĐRMH	Tỷ lệ (%)
A1. Quá trình (Kiểm tra trên lớp, bài tập, đồ án, ...)		0%
A2. Giữa kỳ	G1, G2, G3, G4	30%
A3. Thực hành		0%
A4. Cuối kỳ	G1, G2, G3, G4, G5, G6	70%

## 7. QUY ĐỊNH CỦA MÔN HỌC (Course requirements and expectations)

- Quy định về giờ giấc, chuyên cần, kỷ luật trong khóa học: sinh viên đến lớp theo quy định chung của trường.
- Quy định liên quan đến các sự cố có thể xảy ra với bài thi, bài tập: theo quy định chung của trường.
- Quy định sử dụng phương tiện học tập: theo quy định của giáo viên.
- Nghe giảng trên lớp, nghiên cứu tài liệu ở nhà, tích cực tham gia thảo luận.
- Sinh viên phải làm các bài tập, phải chuẩn bị trả lời các câu hỏi và đọc trước những tài liệu theo yêu cầu của giảng viên trước mỗi buổi học.

## 8. TÀI LIỆU HỌC TẬP, THAM KHẢO

### **Giáo trình**

1. Mount, David W (2001). *Bioinformatics: Sequence and Genome Analysis*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. ISBN: 9780879695972.
2. Nguyễn Văn Cách (2005). *Tin sinh học*. Nxb. Khoa học Kỹ thuật.

### **Tài liệu tham khảo**

1. Watson, James, Tania Baker, Stephen Bell, Alexander Gann, Michael Levine, and Richard Losick (2003). *Molecular Biology of the Gene*. 5th ed. San Francisco, CA: Benjamin Cummings. ISBN: 9780805346350.
2. Trần Linh Thuộc (2004). *Thực tập Bioinformatics*. Đại học quốc gia Thành phố HCM.
3. Chu Văn Mẫn, Đào Hữu Hồ (2001). *Giáo trình thống kê sinh học*. NXB Khoa học kỹ thuật.

## 9. PHẦN MỀM HAY CÔNG CỤ HỖ TRỢ THỰC HÀNH

*Tp.HCM, ngày 03 tháng 03 năm 2016*

**Trưởng khoa/bộ môn**

(Ký và ghi rõ họ tên)

**Giảng viên biên soạn**

(Ký và ghi rõ họ tên)

TS. Nguyễn Gia Tuấn Anh

ThS. Nguyễn Thành Trung