**生物信息学综合实验课程教学大纲**

Course Outline

|  |
| --- |
| 课程基本信息（Course Information） |
| 课程代码（Course Code） | BI390 | 学时（Credit Hours） | 64 | 学分（Credits） | 2 |
| 课程名称（Course Name） | （中文）生物信息学综合实验 |
| （英文）Bioinformatics workshop |
| 课程性质(Course Type) | 培养计划课程Compulsory |
| 授课对象（Target Audience） |  |
| 授课语言(Language of Instruction) | 中英双语Chinese/English |
| 开课院系（School） | 生命科学技术学院生物信息学与生物统计学系Department of Bioinformatics and Biostatistics, School of Life Science |
| 先修课程（Prerequisite） | 生物信息学、生物化学、分子生物学、程序设计、数理统计、生物统计学等Introduction to Bioinformatics, Programming, Statistics |
| 授课教师（Teacher） | 吴茂英Maoying, Wu | 课程网址(Course Webpage) | http://cbb.sjtu.edu.cn/course/bio2407 |
| \*课程简介（Description） | 本课程是对前面三年生物信息学本科专业学习的一个总结，强调学生对核酸、蛋白质序列的分析算法、结构和功能注释方法的认识和实现，同时还对组学数据分析中常用的统计学、信息学、机器学习和数据挖掘方法进行全面总结，并结合学科发展的前沿，对出现的新技术和新方法进行深入的分析和探索，旨在培养学生全面的生物医学数据科学研究和开发的能力。 |
| \*课程简介（Description） | This course is comprised of a series of computational hands-on exercises emphasizing the fundamentals of computational algorithms on nucleotides, protein sequences and the respective functional annotations; it also includes an introduction to the analysis of network-based complex biological systems. Topics covered in the course include principles, methods, and tools applied for sequence alignment, motif discovery, network modeling and analysis, as well as recently emerging research areas - such as next generation sequencing data analysis and deep learning. This course is designed for high-level undergraduates and graduates with strong background in both biology and computer science. Basic programming skills in C/Perl/Python programming are needed. |
| 课程教学大纲（course syllabus） |
| \*学习目标(Learning Outcomes) | 1．掌握生物信息学中常用的数据分析工具，并熟练运用进行生物医学数据分析。2．掌握生物信息学各个领域的核心算法，并能用掌握的计算机语言实现。 3．能根据提出的生物学问题开发算法，并加以实现，有效解决问题4. 了解学科各个领域最新的发展动向，对学科的发展趋势有所认识 |
| \*教学内容、进度安排及要求(Class Schedule& Requirements) |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 教学内容 | 学时 | 教学方式 | 作业及要求 | 基本要求 | 考查方式 |
| 生物信息学中的高性能计算 | 8 | 讲解/上机实验/案例实现 | 掌握并行集群构建的方法，MPI并行编程的实现，GPU编程实现（可选）；了解云计算平台 | 掌握并行计算的分类，并能加以实现， | 完成编程作业 |
| 基因组序列分析方法 | 8 | 讲解/上机实验 | 完成比较基因组分析，预测基因组中的各种组件 | 掌握比较基因组分析的各种策略；对于motif发现的各种算法有深入的认识 | 上机作业/算法开发 |
| 二代测序数据的分析 | 8 | 讲解/上机实验 | 各种二代测序平台数据的处理和分析，尤其是各种数据格式的解读和深入挖掘 | 了解各种二代测序技术的关键要点以及数据处理方法；掌握各种短序列比对、拼接等不同策略和算法 | 上机作业/工具开发 |
| 组学数据分析与统计 | 12 | 讲解/上机实验 | 深度重测序数据解读并找到各种突变；RNAseq数据的分析发现差异表达基因；元基因组数据的解读和分析；表观基因组数据的分析 | 能运用各种工具和开发统计学分析工具，结合各种算法对组学数据进行挖掘 | 上机作业 |
| 生物医学数据分析中的数据挖掘和机器学习方法 | 8 | 讲解/上机实验 | 生物医学大数据的预处理；无监督学习和有监督学习的原理、分类和应用，以及在数学底层的联系；深度神经网络和深度SVM在生物医学数据分析中的应用 | 掌握数据挖掘的各种方法，能在不同的计算平台上灵活运用这些方法对生物医学数据进行分析 | 上机作业 |
| 生物复杂网络的建模与分析 | 8 | 讲解/上机实验 | 基因调控网络、蛋白相互作用网络、代谢网络模型的构建和结合，以及网络模块的发现及其应用 | 基于各种组学数据重构静态和动态网络，并采用网络发现某些生物学过程的关键要素和作用模式 | 上机作业 |
| 统计遗传学分析 | 4 | 讲解/上机实验 | 多种基于家系和基于人群的分析方法和结果的评估；全基因组关联分析；Meta分析中的固定效应模型和随机效应模型 | 掌握统计遗传分析从实验设计到样本收集和数据分析的手段 | 上机作业 |
| 结合学科前沿可选项目开发和报告 | 8 | 学生口头报告 | 结合学科最高水平杂志文章和自己的研究思路，开展项目研究 | 独立从事生物信息学的研究和开发的能力 | 口头报告 |

 |
| \*考核方式(Grading) | 项目报告（70%） + 口头报告 （30%）Project reports (70%), Oral presentation (30%) |
| \*教材或参考资料(Textbooks & Other Materials) | 无参考书，采用经典的文章和新近发表的高水平杂志文献等。No Textbook. |
| 其它（More） | 无 |
| 备注（Notes） | 无 |

备注说明：

1．带\*内容为必填项。

2．课程简介字数为300-500字；课程大纲以表述清楚教学安排为宜，字数不限。