课程大纲
COURSE SYLLABUS

1.	**课程代码/名称** **Course Code/Title**	**ESE5013 生物信息学在环境科学中的应用/Applied Bioinformatics in Environmental Science**
2.	**课程性质** **Compulsory/Elective**	专业选修课/Elective
3.	**课程学分/学时** **Course Credit/Hours**	3
4.	**授课语言** **Teaching Language**	中英文/Chinese and English mixed
5.	**授课教师** **Instructor(s)**	夏雨 XIA Yu
6.	**先修要求** **Pre-requisites**	
7.	**教学目标** **Course Objectives**	
	本课程的教学目的包括： 1) 掌握高通量测序平台的测序原理 2) 掌握生物信息学常用工具的算法原理（BLAST, HMM, BWT） 3) 掌握基本生物信息学工具的使用 4) 掌握基本 Linux 命令行操作及编程 The course is designed to help students: 1) Understand the sequencing principles of high-throughput sequencing platforms. 2) Understand the fundamental algorithms in bioinformatics（BLAST, HMM, BWT）. 3) Master basic bioinformatic data mining of Metagenomic dataset. 4) Master basic Linux command line and script writing.	
8.	**教学方法** **Teaching Methods**	
	课堂授课结合上机实践	
9.	**教学内容** **Course Contents**	
	Section 1	**绪论 （2 学时）**/ Introduction (2 class hours)
	Section 2	**第一部分 生物信息学算法导论 （10 学时）**/ Part I: Fundamental Knowledge of bioinformatic Algorithms (10 class hours in total) 1. Principle of sequencing platforms 2. Alignment algorithms（BLAST, HMM, BWT） 3. Introduction of bioinformatics resources
	Section 3	**第二部分 宏基因组学数据分析演示（20 学时）**/ Part 2 : Metagenomic data analysis demo (20 class hours) 1. Data basics 2. Basic QC 3. Community Analysis 4. Metagenomic assembly

		5. Metagenomic annotation 6. Metagenomic binning analysis
	Section 4	**第三部分宏基因组学数据分析实战（10 学时）/** Part3 Student Metagenomic Projects （10 class hours） 1. Community Analysis 2. Metagenomic assembly 3. Metagenomic annotation 4. Metagenomic binning analysis
	Section 5	
	Section 6	
	Section 7	
	Section 8	
	Section 9	
	Section 10	
	……	另有 6 学时用于学生演讲和课堂讨论，总设计 48 学时 The remaining 6 class hours are used for student presentation, in-class discussion. The total number of hours is 48
10.	**课程考核** **Course Assessment**	
	请再此注明：①考查；②分数构成。 Attendance 10% Class Performance 10% Presentation 20% Final project 60%	
11.	**教材及其它参考资料** **Textbook and Supplementary Readings**	
	主要参考教材/Major textbooks： 1) 《鸟哥的 Linux 私房菜》，第三部分：学习 Shell 与 Shell scripts，http://cn.linux.vbird.org/linux_basic/linux_basic.php#part3） 2) 《生物信息学生 R 入门教程》，糗世界，http://qiubio.com/archives/3740 3) 《The Biostar Handbook》，István Albert，https://www.biostarhandbook.com/	

课程简介：

近十年来高速发展的高通量测序技术极大促进了宏基因组学方法在环境微生物研究领域的应用，宏基因组学该方法正在逐渐成为表征环境微生物的生态功能、识别微生物互作机制乃至判定微生物遗传进化关系的主流研究手段。针对这一新兴领域、本课程将以课堂讲授、课堂演示操作与数据分析实战相结合的方式，培养学生在高通量测序及宏基因组学研究领域的理论基础和实际上机操作能力。课程内容涵盖以高通量测序为基础的宏基因组学研究中所涉及的 1）高通量测序原理、2）数据比对算法 BLAST,HMM 以及 BWA 的核心原理、3）Linux 命令行编程基础；以及 4）最先进水平的宏基因组数据分析工具的使用方法及

算法原理。最后学生将开展一组自主选题的宏基因组学数据分析工作，在授课老师和助教的指导下解决实际数据处理中会遇到的计算机编程问题以及知识挖掘困难，亲身体验宏基因组学研究的工作流程，掌握用宏基因组方法开展科学研究的方法。

Course Introduction:

The rapid development of high-throughput sequencing technology in recent ten years has boosted the widespread application of metagenomics in environmental microbiome investigations. Given the robustness, metagenomics is becoming a standard culture-independent method to characterize phylogenetic affiliation and reveal ecological functions of environmental microbes. What accompanies the continuous upgrading of bioinformatics tools is the gradual construction of knowledge hierarchy in high throughput sequencing-based metagenomics. Thereafter, this course is designed to let the students build knowledge basis on sequencing and algorithms principles involved in modern metagenomics and master basic analytical skills to handle state-of-the-art metagenomic data mining. To establish practical experience in programming debugging and problem-solving, students will accomplish the course by practicing what they have learned in solving a self-selected scientific question by real metagenomic analysis in a one-month-long project. This course will be a perfect suit for graduates students whose research involves the application of high-throughput metagenomic sequencing.