# PhD in Molecular Biology and Biotechnology (MBBT) (Programme Code: PHDBT)

# Program Educational Objectives (PEOs), Program Outcomes (POs) and Learning Outcomes (LOs)/ Course Outcomes (COs)

January, 2021



# Department of Molecular Biology and Biotechnology

Tezpur University Tezpur, Assam, India

## Preamble

PhD in Molecular Biology and Biotechnology offer by the department provides opportunity to carry out advance research in diverse areas of modern biology. It also contains a course component of 16 credits aimed at imparting knowledge on methodologies, ethical guidelines and critical analysis of research articles.

#### 1. Introduction

PhD in Molecular Biology and Biotechnology (MBBT) is an interdisciplinary research programme for developing skills to carry out cutting edge research in modern biology. The course will enable the scholars to develop confidence for taking up biological problems independently, initiate bio-entrepreneurship, contribute to academia and/or in industry.

#### 2. Qualification descriptors for the scholars

Knowledge and Understanding

- i) In dept knowledge and understanding of recent developments in their research area
- ii)In dept knowledge and understanding of gap in their research domain
- iii) In dept knowledge and understanding to undertake further research in their domain

Skill and Technique

- i) Scholars will be skilled in research proposal.
- ii) Scholars will be skilled in conducting cutting edge research.
- iii) Scholars will be skilled in supervising research activities.

Competence

- i) Scholars will be competent to critically analyse biological problem in their research area.
- ii) Scholars will be able to carry out independent research.
- iii) Scholars will be able to scientifically contribute to their subject area.
- iv) Scholars will have expertise for employment in academia and/or in biotech industries.

#### 3. Scholar Attributes

- i) Scholars with expertise in their research domain.
- ii) Skilled human resource to contribute to the cutting-edge research in modern biology.
- iii) Well trained scholars in the area of their research domain with an ability to contribute to the society through their research findings.

## 4. Program Outcomes (POs)

**PO1**: Scholars will develop research skills.

- PO2: Scholars will be able to critical analyse and understand the biological problems.
- **PO3:** Scholars will be able to design experiments for such biological problems.
- **PO4**: Scholars will gain knowledge and confidence for employment in academic institutions and Biotech industries.
- **PO5**: Scholars will be able to communicate with the common mass regarding scientific benefits.

## 5. Programme structure

Total Credits: 18 Structure of the curriculum

Course category	No of courses	Credits per course	Total Credits
I. Core courses	3	4+4+2	10
II. Elective courses	4	4	4
III. Open Elective	1	4	4
Total credits			18

#### 6. Course structure

Course type	Course title	Lecture (L)	Tutorial (T)	Practical (P)	Contact Hour(CH)	Credits		
Core	BT 710: Research Methodology	2	2	0	4	4		
	BT 711: Review of Literature and Experimental	1	1	2	6	4		
	Techniques							
	RP 799: Research and Publication Ethics	2	0	0	2	2		
Elective	BT 703: Toxinology	3	1	0	4	4		
	BT 716: Omics in Biology	3	1	0	4	4		
	BT 717: Evolutionary Genetics and Immunogenetics		1	0	4	4		
	BT 718: Structural Bioinformatics and Modelling	3	1	0	4	4		
Open	Open Elective		0	0	4	4		
Elective								

#### 7. Mapping: Program Educational Objectives (PEOs) and Program Outcomes (POs)

Course title	P01	P02	PO3	P04	PO5
BT 710: Research Methodology	Х	Х	Х	-	-
BT 711: Review of Literature and	Х	Х	Х	-	-
Experimental Techniques					
RP 799	Х	Х	Х	Х	х
CBCT	Х	Х	Х	-	-
BT 703: Toxinology	Х	Х	Х	Х	Х
BT 716: Omics in Biology	Х	Х	Х		
BT 717: Evolutionary Genetics and	Х	Х	Х	Х	х
Immunogenetics					
BT 718: Structural Bioinformatics	Х	Х	Х	Х	-
and Modelling					
PhD Thesis	Х	Х	Х	Х	Х

## 8. Evaluation plan:

- Understanding of subject is constantly evaluated through discussion and cross questioning with the supervisor.
- Examinations of specific duration (1 Hr/2 Hr) (Course work)
- > Assignment on critical problems related to research subject.
- Constant assessment during research
- > Oral presentations on articles related to research area.
- > Synopsis and progress seminar presentation
- > Research article writing.
- ➤ Thesis writing
- Pre-Thesis presentation
- Evaluation of thesis by external examiner
- > Oral defence of PhD thesis to be evaluated by external examiner
- $\triangleright$
- 9. DETAILED SYLLABUS

#### **BT 710: Research Methodology Course outcome**

CO1: Ability to identify research problems, design experiments and carry out research CO2: Ability to understand ethical concerns, research ethics and biosafety issues. CO3: Ability to analysis of research articles and gain skills in technical writings. CO4: Ability to write research proposals and grant applications. CO5: Ability to use computation tools in their research area.

#### **Course content**

- 1. Introduction to research methodology: Scope
- **2. Methods of Research:** survey, observation, case study, experimental, historical and comparative methods, Difficulties in Biological research Literature review
- **3. Research problem:** defining the research question, identification, selection, formulation of research objectives
- **4. Research design:** Components and Importance, Documentation, presentation and analysis of data: Types of data, Data collection, Methods and tools of data collection, presentation of data, analysis and interpretation of data
- 5. Ethics in research: Institutional ethical committee for human and animal research
- 6. Biosafety: Guidelines for Biosafety, Institutional Biosafety committee and its role
- 7. Plagiarism Research ethics, Pitfall, Software for Plagiarism
- 8. Patents and IPR: Patent laws, process of patenting a research finding, Copy right, Cyber laws
- **9. Research funding**: A brief idea about the funding agencies such as DST, DBT, ICMR, CSIR, UGC, ICAR, International funding
- **10. Technical Writing**: Writing of Research Proposal and Report, Manuscript preparation Quantitative Data Analyses
- **11. Hypothesis testing:** Normal and Binomial, poisson distributions and their property. Tests of significance: Student t test, F- test, Chi square test Correlation and Regression ANOVA –One way and Two way, Multiple range test

## Practical

## **Computer Fundamentals**

Introduction to spread sheet application, features and functions, Using formulas and functions, Data storing, Generating charts/ graph and other features. Statistical data analysis using software like Microsoft Excel, Origin etc

#### **Textbooks:**

- 1. Ranjit Kumar, Research methodology: A step by step guide for beginners, 2<sup>nd</sup> edition, SAGE Publications Ltd., 2005.
- 2. John W. Creswell, Research Design: Qualitative, Quantitative, and mixed methods approaches, 2<sup>nd</sup> edition, SAGE Publications, 2009.

## Suggested Readings

1. Petter Laake, Haakon Breien Benestad and Bjorn Reino Olsen, Research Methodology in the Medical and Biological Sciences,1<sup>st</sup> edition, Academic Press, 2007.

## **BT 711: Review of Literature and Experimental Techniques** Course outcome

CO1: Ability to review literatures and write review of literature in their area of research. CO2: Ability to search journals related to their research area and their publication impacts. CO3: Ability to design methodologies and techniques required for their research.

## **Course content**

1. Review of Literature (CR2)

- a) **Review of Literature:** Data base search for literatures using Pubmed, Web of Science, Google scholar, J-gate, SciFinder, JCCC Database.
- **b)** Writing review of literature:
- c) Citation: Citation Index: Science Citation Index (SCI), h index, I-10 index. Journal Impact Factor (JIF) Features
- d) Seminar presentation

## 2. Experimental Techniques (Laboratory rotation) (CR2) L0-T0-P2-CR2 (A Student need to take ANY TWO of the following experimental techniques courses)

## Molecular Toxinology Laboratory

- 1) Quantitation of Protein: Estimation of protein using Lowry's method
- 2) Electrophoresis: Determination of homogeneity of protein using SDS-PAGE
- 3) Biological characterization of snake venom: Anticoagulant activity, antibacterial, hemolytic, in-vitro tissue damaging, PLA<sub>2</sub> activity, proteolytic activity, etc.

## Innate Immunity and Immunogenetics Laboratory

- 1) Immune Cell Separation and Characterization- separation of PBMC by density gradient, isolation of purified cell populations based on CD antigens
- 2) Characterization of cells by- immunophenotyping( IHC and IF ) and secretory profile (cytokine and chemokines )
- 3) Antigen characterization ELISA and western blot

## Molecular Genetics (Bacteria) Laboratory

- 1) Isolation of spontaneous mutants
- 2) Creation of site specific insertion mutation in bacteria
- 3) Reporter gene fusion study in bacteria

## Human Molecular Genetics Laboratory

- 1) Isolation of genetic material from blood & other tissues by conventional phenol chloroform method and using kits.
- 2) Qualitative and Quantitative assessment of the DNA/RNA: Agarose gel Electrophoresis, PAGE, Colorimetric Assays.
- 3) Polymerase Chain Reaction: Primer designing, PCR amplification reaction for a human gene/ fragment
- 4) Screening for mutations/SNPs in candidate genes by (Single Stranded Conformation Polymorphism (SSCP) and Restriction Fragment Length Polymorphism (RFLP) Analysis.
- 5) Introduction to the databases like OMIM, ENSEMBL, PubMed, etc. and their usage.

## Molecular Endocrinology laboratory

- 1) Reverse transcription polymerase chain reaction (RT-PCR): To detect mRNA expression levels in different cells/tissues or the cells under different experimental regime.
- 2) SDS-PAGE and Western blot: To separate proteins according to their electrophoretic mobility and detect specific proteins in the given sample of cell extract /tissue homogenate by using antibodies.
- 3) Chromatin immunoprecipitation assay (ChIP): To investigate the interaction between proteins and DNA in the cell.

## L2-T0-P0-CR2

## CR=1

CR=1

# CR=1

## CR=1

**CR=1** 

## L2-T0-P2-CR4

#### **Modelling and Simulation Laboratory**

- 1) Computer basic knowledge; hardware, internet browsers, search engines, C & Perl programming languages
- 2) Understanding Biological Databases
- 3) Protein Modeling and Drug design: Homology modeling, model refinement, evaluation of the model, Threading, The drug discovery process, Molecular docking
- 4) Simulation of biomolecular systems: using Gromacs, AMBER

## **Inflammation Laboratory**

- 1) Media preparation: Preparation of different media for cell culture, role of each media ingredients.
- 2) Handling of cells: culture of adherent and non-adherent cells, regular maintenance, how to detect contamination.
- 3) Handling cell culture facilities: Safe practise for using CO2 incubator and Biosafety cabinet, safe disposal technique
- 4) Use of cells for your experimentation: Treatment of cells with stimulus, preparation of lysate for isolation of protein/DNA/RNA

## **Advance Cell Biology Laboratory**

- 1) Fractionation of animal cells and isolation of cell organelles by differential centrifugation.
- 2) Purification of nucleus and estimation of purity using western blot technique.
- 3) Localization of specific cellular proteins in animal cells by immuno-fluorescence microscopy.

## **Microbial Biotechnology Laboratory**

- 1) Biochemical/Microscopic/Molecular methods used to differentiate between archae, eubacteria and eukaryotes.
- 2) Microbial nutrition; different types of culture medium; C/N/P balance and making of culture medium.
- 3) Assay of antibiotics production and demonstration of antibiotic resistance

## **Enzyme Technology Laboratory**

- 1) Isolation and screening of industrially important microorganisms
- 2) Production and purification of extracellular/intracellular enzyme
- 3) Enzyme activity and initial velocity determination
- 4) Enzyme immobilization
- 5) Scale up strategies in enzyme production

CR=1

## **CR=1**

**CR=1** 

CR=1

## **RP 799 Research and Publication Ethics**

Course outcomes:

On completion of this course the students will be able to learn:

CO1: Basics of philosophy of science and ethics, research integrity, publication ethics

CO2: Identification of research misconduct and predatory publications

CO3: Indexing and citation databases, open access publications, research metrics

(citations,

*h-index, Impact Factor, etc.)* 

CO4: Plagiarism tools and their use

CO5: Communication of research finding without ethical violation

## Course Contents:

## THEORY

## RPE 01: PHILOSOPHY AND ETHICS (3 hrs.) [3 Lectures]

- 1. Introduction to philosophy: definition, nature and scope, concept, branches
- 2. Ethics: definition, moral philosophy, nature of moral judgements and reactions

## RPE 02: SCIENTIFICCONDUCT [5 Lectures]

- 1. Ethics with respect to science and research
- 2. Intellectual honesty and research integrity
- 3. Scientific misconducts: Falsification, Fabrication, and Plagiarism (FFP)
- 4. Redundant publications: duplicate and overlapping publications, salami slicing
- 5. Selective reporting and misrepresentation of data

## **RPE 03: PUBLICATION ETHICS [7 Lectures]**

- 1. Publication ethics: definition, introduction and importance
- 2. Best practices / standards setting initiatives and guidelines: COPE, WAME, etc.
- 3. Conflicts of interest

4. Publication misconduct: definition, concept, problems that lead to unethical behavior and vice versa, types

- 5. Violation of publication ethics, authorship and contributorship
- 6. Identification of publication misconduct, complaints and appeals
- 7. Predatory publishers and journals

## PRACTICE

## RPE 04: OPEN ACCESS PUBLISHING [4 Lectures]

- 1. Open access publications and initiatives
- 2. SHERPA/RoME0 online resource to check publisher copyright & self-archiving policies
- 3. Software tool to identify predatory publications developed by SPPU

4. Journal finder / journal suggestion tools viz. JANE, Elsevier Journal Finder, Springer Journal Suggester, etc.

## **RPE 05: PUBLICATION MISCONDUCT [4 Lectures]**

## A. Group Discussions (2 hrs.)

- 1. Subject specific ethical issues, FFP, authorship
- 2. Conflicts of interest
- 3. Complaints and appeals: examples and fraud from India and abroad

## **B.** Software tools (2 hrs.)

Use of plagiarism software like Turnitin, Urkund and other open source software tools

## RPE 06: DATABASES AND RESEARCH METRICS [7 Lectures]

## A. Databases (4 hrs.)

- 1. Indexing databases
- 2. Citation databases: Web of Science, Scopus, etc.

## **B.** Research Metrics (3 hrs.)

1. Impact Factor of journal as per Journal Citation Report, SNIP, SJR, IPP, Cite Score

2. Metrics: h-index, g index, il 0 index, altmetrics

Reference Books:

- 1. A. Bird, Philosophy of Science, Routledge, 2006.
- 2. A. MacIrityre, A Short History of Ethics, London, 1967
- 3. P. Chaddah, Ethics in Competitive Research: Do not get scooped; do not get plagiarized, 2018.ISBN:978-9387480865
- 4. National Academy of Sciences, National Academy of Engineering and Institute of Medicine. On Being a Scientist: A Guide to Responsible Conduct in Research: Third Edition. National Academies Press., 2009.
- D. B. Resnik, What is ethics in research & why is it important. National Institute of Environmental Health Sciences, 1-10, 2011. Retrieved from <u>https://www.niehs.nih.gov/research/resources/bioethics/whatis/index.cfm</u>
- 6. J. Beall, Predatory publishers are corrupting open access. Nature, 489(7415), 2012, 179-179. https://doi.org/10.1038/489179a
- Indian National Science Academy (INSA), Ethics in Science Education, Research and Governance(2019), ISBN:978-81-939482-1-7. http://www.insaindiaxes.in/pdf/Ethics Book.pdf

#### **Elective Courses**

## **BT 703: TOXINOLOGY**

#### Course outcome

CO1: Ability to identify and isolate toxins from natural sources and its impact and benefits. CO2: Ability to review current research trends on toxins. CO3: Ability to evaluate the structure and function of toxins corelation.

#### **Course content**

- 1. **General Introduction:** Definition of toxins, different categories of toxins and venoms, recent trends in venom and toxin research.
- 2. **Bacterial toxins**: Bacterial toxinogenesis, endotoxins, exotoxins, exotoxins, bacterial protein toxins with special reference to cholera, diphtheria and tetanous toxins, molecular mechanism of action of endotoxins, exotoxins, enterotoxins and neurotoxins.
- 3. **Toxins from snake venom.** Snakes and Biological significance of their venoms, composition of snake venom, evolution of venom, 3D structure of some important venom constituents and their mechanism of action (phospholipase A2, cardiotoxin, neurotoxin) three-finger toxins, antivenom and medicinal plants in treatment of snakebite patients.
- 4. Venomics: Proteomics approach to study the venom proteomics.
- 5. **Tools for isolation and characterization of toxins:** Multidimensional chromatographic techniques (gel-filtration, ion-exchange reverse-phase HPLC, SDS-PAGE, 2-dimensional gel electrophoresis), toxin mass fingerprinting, N-terminal peptide sequencing, analysis of protein data by using proteomics softwares.
- 6. **Medicinal and industrial applications of venoms and toxins:** Use of toxin in neurobiology and muscular research, anticancer drug, diagnosis of hemostatic disorders, antibacterial agents, bioinsecticide and other industrial applications.

## Textbooks

- 1. Madagon, Martinks and Parker, Biology of Microorganism,10th edition, Prentice Hall, 2003.
- 2. A.L. Harvey, Snake toxins, First edition, Pergamon, 1991.

## Suggested Readings

- 1. Stephen P. Mackessy, Handbook of Venoms and Toxins of Reptiles, CRC Press, Taylor and Francis Group, 2010.
- 2. Kurt F. Stocker, Medical Use of Snake Venom Proteins, CRC Press, 1990.

#### **BT 716: Omics in Biology**

#### **Course outcome**

- CO1: Ability to understand the fundamentals of genomics and proteomics, transcriptomics and metabolomics.
- CO2: Ability to analyse genome, map genome, evaluate evolutionary process and compare between organisms

CO3: Ability to understand the biological systems using genomics, transcriptomics and proteomics

#### **Course content**

- 1. **Structural organization of genomes**: Introduction to omics in light of the central dogma. Concept of genomics and proteomics (transcriptomics and metabolomics) in the post genome era. Organization of Prokaryotic genome. Organization of organellar genome (mitochondria and chloroplast).Viral genome organization. Organization of eukaryotic genome. Organization and regulation of eukaryotic chromatin structure.Introduction to human genome.
- Genomics: DNA sequencing-principles. Classical vs. Next generation DNA sequencing. Strategies for genome sequencing. Genetic and physical mapping of genome. RFLP, SSLP, STSs, SNPs and other markers. Prokaryotic (H. influenzae) and eukaryotic (human) genome projects: case study. Gene annotation. Comparative genomics. Microarrays. Accessing and retrieving genome project information from web. Metagenomics.
- 3. **Transcriptomics:** Significance of transcriptomics, Assembly, alternative splicing, RNA editing, important file formats and techniques in Transcriptomics, Sequencing based approaches to study transcriptomes, Gene expression profiling, Non coding RNA and miRNA: discovery and analysis.
- 4. **Functional Genomics and Pharmacogenetics:** Introduction to pharmacology. Screening strategies for target gene for diseases and drug action. High throughput screening. Genetic variability and drug action. Pharmacogenomics in drug development.
- 5. **Proteomics:** Significance of proteomics. Analysis of amino-acid composition of proteins and protein structure function. Sequencing of proteins: N-terminal sequencing. Mass spectrometry and proteomics. Protein separation techniques in proteomics: 2D PAGE and HPLC. Peptide mass fingerprinting, post translational modification. Protein-protein interaction analysis by yeast two-hybrid system. Interactome.
- 6. **Metabolomics:** Significance of metabolomics. Metabolic pathways. Different techniques and approaches in metabolomics. NMR and Mass Spectrometry. Targeted and non-targeted metabolomics.
- 7. **Structural genomics:** Determining gene function by sequence analysis. Conserved protein structures and gene function.

#### Textbooks

- 1. T.A. Brown, Genomes 3, 3 rd edition, Garland Science, 2006.
- 2. Primrose B. and Twyman, R. Principles of Genome Analysis and Genomics, 7th edition. Blackwell, 2006.

#### **BT 717: Evolutionary Genetics and Immunogenetics**

#### **Course outcome**

CO1: Ability to understand the theories of classical and modern evolution CO2: Ability to analyse molecular population genetics as well as other related areas of genetics CO3: Ability to analyse human migration and its impact CO4: Ability to co-relate evolution and immunity

#### **Course content**

- 1. Concept and theories of evolution (classical to modern); concept of species and modes of speciation: sympatry, allopatry, stasipatry & parapatry; mechanism of speciation; isolation mechanisms, non-random and random breeding: inbreeding and assortative mating; path diagram construction and inbreeding construction and inbreeding co-efficient, allelic identities by descent; heterosis and heterozygous superiority
- 2. Molecular population genetics: molecular evolution (neutral theory, punctuated equilibrium); molecular clock, molecular evolution and phylogenetic tree; DNA-DNA hybridization; restriction enzyme sites; nucleotide sequence comparison and homologies; human phylogeny: hominid evolution: anatomical, geographical, cultural, molecular phylogenetics of Homo sapiens.
- 3. Admixture: meeting of human population and its genetic imprint; detection of admixture (based on allele frequencies & DNA data); Y chromosome and & mitochondrial DNA markers in genealogical studies; peopling of continents (Europe, Africa, Asia): geogenomics and human migration; culture and human evolution:
- 4. Learning, society and culture; relative rates of cultural and biological evolution; social Darwinism; sociobiology & economics of genetics (econogenetics).
- 5. Human Migration- out of Africa migration and the genetic evidence, shaping of human genome by pathogen pressure and selection
- 6. TLRs and ILRs evolutionary and clinical genetics perspective Natrual Killer Cell receptors and their MHC class I ligands- evolution and immunity

#### Textbooks

- 1. R. H. Tamarin, Principles of Genetics, 6<sup>th</sup> edition, William C Brown Pub ,1998.
- Daniel L. Hartl and Elizabeth W Jones , Genetics: Principles and Analysis, 4<sup>th</sup> edition, Jones & Bartlett Pub, 1997.
- 3. Sylvie Lesage , Immunogenetics: Tolerance and Autoimmunity, Nova Science Publishers, 2010.

## **Suggested Readings**

- 1. W. H. Li, Molecular Evolution, illustrated edition, Sinauer Associates, Incorporated, 1997.
- 2. Motoo Kimura , The neutral theory of molecular evolution, Reprint edition, Cambridge University Press, 1985.

## BT 718: Structural Bioinformatics and Modelling

## **Course outcome**

CO1: Ability to determine the macromolecular structure using computational tools CO2: Ability to perform databases search and review the literature CO3: Ability to perform molecular modelling and understand their structure CO4: Ability to perform molecular dynamic simulation methods

## **Course content**

## 1. Introduction to Macromolecular Structure:

Proteins: primary, secondary, tertiary and quaternary structure; Nucleic acids: DNA-A, B, and Z forms, RNA, Molecular Viewers: Rasmol, Chimera, Pymol, and VMD.

## 2. Bioinformatics databases:

Introduction, nucleotide sequence databases, protein sequence databases, sequence motif databases, protein structure databases, other databases: enzyme, pathway, disease, literature and specialized databases.

## 3. Sequence Alignments:

Introduction, concept of alignment, scoring matrices, PAM, BLOSUM, alignment of pairs of sequences, alignment algorithms, Heuristic methods, multiple sequence alignment.

## 4. Protein Modelling and Drug Design:

Methods of Protein modeling, homology or comparative modeling, model refinement, evaluation of the model, Threading or fold recognition, Ab initio/De novo method. The Drug discovery process, SAR and QSAR techniques in Drug design, Graph theory, Molecular docking.

## 5. Introduction to force fields and modeling:

Introduction to force fields, general features of molecular mechanics force fields; bond stretching; angle bending; torsional terms; out-of-plane bending motions; non-bonded interactions; electrostatic interactions; van der Waals interactions; hydrogen bonding.Energy minimization; non-derivative minimization methods, the simplex method & the sequential univariate method; first-order minimization methods; second derivative methods, the Newton-Raphson method; selection of minimization method.

## 6. Molecular Dynamics Simulation Methods

Molecular dynamics using simple methods; molecular dynamics with continuous potentials; setting up and running a molecular dynamics simulation; constraint dynamics; time-dependent properties; incorporating solvent effects into molecular dynamics; conformational changes from molecular dynamics simulations.

## **Textbooks:**

- 1. Jenny Gu, Philip E. Bourne, Structural Bioinformatics, 2<sup>nd</sup> edition, Wiley-Blackwell, 2009.
- 2. Zhumur Ghosh and Bibekan and Mallick, Bioinformatics: principles and Application, 1<sup>st</sup> edition, Oxford University Press, 2008.
- 3. Andrew R. Leach. Molecular Modelling, Principles & Applications, 2<sup>nd</sup> edition, Prentice Hall, 2001.

## **Suggested Readings**

- 1. Carl Branden and John Tooze, Introduction to Protein Structure, 2<sup>nd</sup> edition, Garland publishing Inc,1999.
- 2. David Mount, Bioinformatics Sequence and Genome Analysis,2nd edition, Cold Spring Harbor Laboratory Press, New York, 2004.