

## Article Information

Received date: May 16, 2016

Accepted date: May 23, 2016

Published date: May 26, 2016

## \*Corresponding author

Mgavi E Brathwaite, New York  
University Tandon School of  
Engineering, Six Metro Tech  
Center, United States,

Tel: 646-997-3950;

Email: m.brathwaite@nyu.edu

Distributed under Creative Commons  
CC-BY 4.0

## Editorial

## Online Bioinformatics Education

Mgavi E Brathwaite

*New York University Tandon School of Engineering, Six Metro Tech Center, United States*

## Introduction

Incorporating digital and online tools to increase the options and flexibility in education is a rapidly growing area of pedagogy. Students can collaborate on projects without being in the same vicinity and download archived live lectures for further review, or access recorded lectures in case they missed it. Students can be anonymously polled about topics during lecture to ascertain class comprehension of subject matter. Timed quizzes as well as proctored exams at secure locations are all part of the online experience.

With the growth of sequencing technology, an exponential amount of genomic data is continuing to be generated. This has increased the demand for highly trained scientist in the discipline of Bioinformatics to make sense of this data. Considering that Bioinformatics is quantitative and an In Silico discipline makes it a great match for online education.

## Methods and Materials

## Program Design

A program should be based around a core set of skills. Since Bioinformatics is an interdisciplinary field that combines genomics, computer science, math, probability and statistics the courses should reflect this dynamic.

The core course for a foundation in Bioinformatics has to cover the following.

1. Nucleotide and Protein Sequence Analysis
2. Common Algorithms in Bioinformatics
3. Application of Statistical Methods in Biology
4. Programming in the form of scripting (Perl, Python, and R).

A 30 credit, at minimum, Masters Program should consist of four core courses and six electives. The program should be flexible to accommodate students from two different skill categories:

1. Life Science
2. Non-Life Science

Hence, there should be two tracks. The life science students will need to develop computational skills to be successful in the program. This includes:

1. Mathematics and Statistics for quantitative biological analysis. Preferably a course from within the program that covers Discrete Math, Multivariate & Bayesian Statistics, as well as Probability.
2. Introduction to computer programming, including UNIX. This prepares the life science student to operate from the command line. The student will learn UNIX commands, Shell Scripting, a computer programming language such as Python, and how to apply these tools to operating and interacting with genome browsers, as well as creating analysis pipelines.

The Non-life science student traditionally will have a background in programming, although this may not include scripting languages such as Perl, Python and R. They will also have had the requisite math to understand the quantitative theory and applications of genomic big data analysis.

What they often lack is a fundamental understanding of genomics. Their life science education may not extend beyond undergraduate biology, chemistry, as well as cellular and molecular biology. To provide a solid foundation in genomics a course designed to bridge their understanding the data that they will be analyzing is necessary. This course would include virtual labs to demonstrate key concepts:

1. Basic introduction to molecular biology.
2. Gene Cloning.
3. Genomic and cDNA libraries

4. Polymerase Chain Reaction.
5. Sequencing.
6. Analysis of Gene Expression.
7. Genomic Analysis.
8. Analysis of Genetic Variation.
9. Post-Genomic Analysis.
10. Transgenic Organisms.

The remaining electives should consist of:

- A. Transcriptomics.
- B. Proteomics.
- C. Quantitative Population Genetics.
- D. Cancer Genomics.
- E. Systems Biology.
- F. Cloud & High Performance Computing.
- G. Guided Studies I.
- H. Guided Studies II.
- I. Genomic Web Applications.
- J. Special Topics.

Transcriptomics is taught in R with the Bioconductor library. Courses B, C, D, & E incorporate R and Bioconductor as the primary tool for data analysis. Cloud and high performance computing will introduce the student to web services that provide a platform for Big Data analysis. This includes services such as Microsoft Azzure, Amazon Web Services, and Google Cloud Platform.

The Guided Studies courses are created to provide the student an avenue to apply their skills in a real world environment. This is an all-around win-win for all parties involved. The student gets the opportunity to demonstrate their Bioinformatics skills and acumen, which often leads to publication, job offers and a solid reference. The Principle Investigator or Lab Head gets some free Bioinformatics help and the academic program gets recognition for excellence in Bioinformatics education.

This prepares the recent graduate for the job interview question about their experience, at which point they can show a body of work to verify their acumen.

A program designed along the lines that I have described can be successful in the new digital age of education.