Dissecting the Tools of Bioinformatics: Building SpelFinder, a primer identification tool in Python

Charles Hardnett and Cynthia Bauerle
Departments of Computer Science and Biology, Spelman College, Atlanta, GA 30314

OBJECTIVE
To help students better understand the construction and utility of standard sequence analysis tools

ABSTRACT
As participant in the National Genomics Research Initiative supported by HHMI, the Biology department at Spelman College offered a course-based research opportunity for first year science students to isolate and characterize novel mycobacteriophages. In fall 2008, twenty first year students isolated novel mycopathogens and conducted standard microscopic and molecular characterization. In spring 2009, fourteen of these students, joined by six advanced biology majors, conducted genomic annotation of one phage isolate using standard sequence analysis approaches and bioinformatic tools available via the Internet. While prior experience applying bioinformatics tools varied among the cohort, no students described any prior experience in computer programming. To help students better understand the construction and utility of standard molecular analysis tools, we developed a course activity in which students used the Python programming language to build a primer identification tool that they named SpelFinder. Students were introduced to basic elements of the Python language in a hands-on group exercise, and then used these elements to build simple algorithms. Ultimately, students assembled algorithms into a script that analyzes an input nucleotide sequence and reports potential primer target sites as output. The program was based on an algorithm that reflected four relevant biological criteria. Writing the program de novo in Python introduced students to basic programming strategies for developing functional algorithms useful in standard bioinformatics tools. Specifically, students were challenged to understand how biological criteria could be translated into a set of rules that drives the algorithm of a sequence analysis software program.

PYTHON BASICS
Python is a general purpose, highly readable, high level computer programming language.

Python program may be created in a standard text editor such as Notepad or TextEdit. Once saved, the program may be run by using the python command to execute the program. The output will appear on the next line:

```
``` PYTHON BASICS
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```