|  |  |
| --- | --- |
| 1. **TOPIC** | Manipulating biological sequences using BioPython for biologists |
| 1. **WHO IS IT FOR AND WHAT DO THEY KNOW?** | Researches that are interested in learning to use programmatic tools to work with biological sequences (DNA, RNA, and amino acids). They should have at least a beginner-level understanding of a programming language and ideally have a basic understanding of strings and lists. All terms will be covered prior to their usage, however. |
| 1. **WHAT WILL THEY LEARN?** | Attendees will learn how to leverage BioPython to examine and manipulate DNA/amino acid sequences found in FASTA files. |
| 1. **3-6 CONCEPTS** | 1. Strings in Python    1. Splitting, subsetting, transforming, and storing 2. How BioPython uses strings (very similar to base Python!)    1. Seq object:       1. BioPython-specific functions like translate       2. Splitting, subsetting, transforming, and storing 3. Reading a FASTA file using BioPython    1. Read in a FASTA file:       1. Translate sequence          1. Write to new FASTA file       2. subset sequence          1. Write to new FASTA file 4. Bonus:    1. Change the headers of sequences in a FASTA file and write to a new FASTA file. |
| 1. **DATA OR TASK/GOAL** | FASTA file with one sequence  FASTA file with ten sequences |
| 1. **INTRO MATERIAL OUTLINE** | Big picture: why do we need programming to analyze sequences? A couple of examples so that people can see that this is something worth their time and attention.  More detailed and one slide per number:  1.Strings (and maybe lists?) in Python and what you can do with them  2. The concept of BioPython and why it was created  3. What variables, objects, attributes, and functions are (in very, very simple terms)  4. What we will have learned by the end of the lesson. |
| 1. **CONCEPTS, IDEAS, EXCEPTIONS, USE CASES YOU WON’T COVER** | Won’t cover BLAST, alignment, motif searching, GENBANK files, ALN files, internal alignment functions. This is just the basics to get people interested in BioPython, with enough info to get them started on their journey toward programmatic analyses of sequences. |
| 1. **WORKSHOP DESCRIPTION AND PREREQUISITES** | Want to use programming to efficiently explore, modify, and write biological sequence (DNA, RNA, and amino acid) data in your research? Learn the necessary foundational skills using BioPython! In this workshop, we will learn how to do common and useful tasks that allow for modification of hundreds of sequences in seconds. This includes reading in FASTA files, editing the names of sequences, and writing modified sequences to new files. Attendees will hopefully have a beginner-level understanding of Python, like knowing what strings, lists, and variables are. *Recordings will be available for those who cannot attend every lesson. Please register for the workshop to get access to the recordings (active Northwestern NetID required).* Prerequisites: Python Fundamentals bootcamp or similar beginning Python tutorial or workshop. |
| 1. **INSTALLATION/ PREPARATION INSTRUCTIONS** | Online instructions:  Google colab instructions: |