Syllabus

The NASA 2020 JPL-HBCU Internship Project will take place virtually, but be hosted primarily out of NASA's Jet Propulsion Laboratory in Pasadena, CA. This a 5-week long internship project comprised of 10 undergraduate students and 2 faculty members from two Historically Black Colleges and Universities (HBCUs), Howard University and Tuskegee University.

The participating students will be computer science majors. The 5-week internship/workshop will entail introductory lessons and assignments in general bioinformatics, followed by group research projects focused on assessing various metagenomic, read-based taxonomic classification and functional assignment pipelines and contamination removal, and then applying the dominant method to 20+ environmental metagenomes from NASA JPL spacecraft-assembly rooms. Though contingent upon the progress made during the 5-week program, publications with the students are anticipated to result (as has been the case with previous allocations granted from XSEDE/Jetstream, e.g. https://portal.xsede.org/web/xup/publications/-/publications/29856, thanks to allocation TG-MCB200008 in that case).

I have in the past been fortunate enough to utilize XSEDE/Jetstream resources to successfully handle larger, though shorter, workshops (including the ~2-week STAMPS course at MBL in Woods Hole MA, USA, and the ~2-week DIBSI course at UCDavis in CA, USA, and our prior 2-day JPL-SAGAn workshop). With how wonderfully easy that made those workshops, I was hoping we would be able to provide this infrastructure for the students from HBCUs for the duration of this 5-week long internship project (and a little longer to help wrap anything up).

Week 1 (beginning July 13, 2020)
- Introductory lectures and assignments for general bioinformatic concepts, e.g.:
    - Common sequencing platforms/how we sequence nucleic acids
    - Amplicon sequencing vs metagenomics (https://astrobiomike.github.io/misc/amplicon_and_metagen)
    - Read-based metagenomic approaches vs assembly approaches
    - Assembly (https://astrobiomike.github.io/genomics/de_novo_assembly)
    - Recovering genomes from metagenomes (https://astrobiomike.github.io/genomics/metagen_anvio)
    - Phylogenomics (https://astrobiomike.github.io/genomics/phylogenomics)

Weeks 2 and 3 (beginning July 20, 2020)
- Focusing on read-based taxonomic classification of metagenomic data
    - Designing and implementing tests and validations of multiple available read-based taxonomic assignment tools/pipelines and contamination removal
        - e.g.: kraken2 -> bracken; ganon; diamond -> MEGAN
    - Applying the decided approach to our 20+ metagenomes from NASA JPL spacecraft-assembly rooms

Weeks 4 and 5 (beginning August 3, 2020)
- Focusing on read-based functional annotation of metagenomic data
    - Designing and implementing tests and validations of multiple available read-based functional annotation tools/pipelines and contamination removal
        - e.g.: humann2; mi-faser; seqscreen
    - Applying the decided approach to our 20+ metagenomes from NASA JPL spacecraft-assembly rooms
    - Collating all results and working on writing up any potential manuscripts

Thank you for your consideration.