What is Bioinformatics? It's a BLAST

What is lurking in the soil? Utilizing DNA databases to identify pathogenic organisms in soybean fields

How can we identify individual organisms in a complex mixture of DNA? What computer tools are needed to decode the source of a DNA sequence?

Background

This lesson illustrates what scientists do after they have completed DNA extractions by introducing basic DNA sequence analysis. Importantly, this lesson goes beyond the visual appeal and celebrity status of the commonly-deployed strawberry DNA extraction by introducing students to DNA analyses through the use of bioinformatics.

Bioinformatics is a combination of computer science and biology that aims to organize, analyze, and interpret biological data. With the success of biotechnology in DNA sequencing, biologists need to use computer science to handle such large, rich databases from organisms being studied. DNA is both a unifying and defining characteristic of all living things. Even in its most raw form, the sequence of As, Ts, Cs, and Gs of DNA is a powerful tool of discovery. Bioinformatics can provide new insights in biology. For example, comparisons of DNA sequences across living organisms have helped biologists describe new species, discover new genes, and predict the health of populations. In agriculture, decoding DNA has boosted our food supply by identifying, tracking, and incorporating genes that provide protection from pathogens and pests. Likewise, cracking the DNA code has helped classify and better detect emerging pathogens that cause disease and substantial crop losses.

Disease accounts for up to 20% of the world’s annual food losses, making pathogens a top threat to food crops. Many pathogens inhabit agricultural fields and attack food crops during the growing season. Soil contains a diversity of microbes, which may include fungi, oomycetes, animals, protists, and bacteria. The community of microbes can be examined by extracting DNA directly from soil and decaying plant matter (residue) found within a farmer’s field. Importantly, not all microbes in soil are pathogens; some are beneficial to growing plants (i.e., nitrogen-fixing bacteria) or help to decompose dead plants (saprophytic bacteria). Following DNA extractions, there are DNA sequencing techniques that can be used to survey some of the more common groups of microbes, including bacteria and fungi. The sequencing techniques examine regions of the genome that are generally different among closely-related species. These regions are called **barcodes**, defined similarly to the barcodes used in supermarkets to identify unique products with a scanner at checkout.

The DNA barcodes are different between eukaryotes (fungi) and prokaryotes (bacteria). In the scenario below, DNA sequences of the barcoding region were used to survey the fungal and bacterial diversity of a field that has given a farmer trouble when soybeans are grown. Specifically, soybean plants continue to produce low seed counts (also termed ‘low yield)’ with noticeable symptoms for disease on the plants (stems and leaves). Multiple applications of a recommended fungicide treatment have not slowed the spread of disease, and more soybean plants are beginning to show signs of disease.

Generally, the sequence of determining an organism is: 1) soil sample collection, 2) DNA extraction, 3) amplification via PCR (targeted or barcode markers), 4) sequencing (barcode markers) or genotyping (targeted markers), 5) analysis. Essentially, this lesson is an introduction to the popular, public database managed by NCBI that looks at sequence analysis, the last operation listed above. The DNA sequence given below can be compared to the database by using the BLAST function. The results from the BLAST will give alignments (i.e., side-by side-comparisons) of the sequence you submit (query) and hits (subject) found in the NCBI database. By clicking on the top hits, the DNA sequences submitted can be identified by species, gene, or genetic marker. Each DNA sequence in the database has its own page that gives additional information, such as the scientists that uploaded the sequences to the database. There are other parameters displayed following a BLAST, which could be further explored (see NCBI presentation deck). Students could be asked to write up a short essay of one of the pathogens and how it could be managed to limit its impact on food crops.

Scenario

Shadybrook Farm is having trouble with production of soybeans and has noticed increasing levels of disease, despite the use of fungicides. The farm manager elects to send a soil sample to get tested to a diagnostic laboratory to try to identify potential disease-causing pathogens. The laboratory produced a report that contained the top five most common DNA sequences from the soil. Your job as a bioinformatician is to help the farm manager identify the species, using the given sequences of DNA, through the use of public databases and basic bioinformatic tools.

You will be given a DNA sequence from a processed soil sample collected from Shadybrook Farm. Identify this organism, its classification, and how it interacts or affects soybeans.

Materials

computer with access to internet
sequences from grownextgen.org/go/blast

Procedure

1. Search the NCBI database to identify one of the sequences of DNA using the following steps:
   a. Visit: https://www.ncbi.nlm.nih.gov
   b. Click on ‘BLAST’ in the right column.
   c. Click on ‘Nucleotide blast’; paste in the DNA sequence you’ve been assigned in the rectangle; submit; wait for result.

Reflection

1. Which organism does this sequence identify?
2. What is its classification (bacteria, fungi, nematode, etc)?
3. Is it a pathogen—does it cause a disease that affects soybeans? If so, what disease, and how does it affect the plant?

4. What do the experts recommend for management of this pathogen?

5. Can you think of other organisms and their DNA that may be found in the soil sample?

Collaborate with others in your group to determine which of the organisms could be causing the trouble on Shadybrook Farm. Create a plan to help the farmer address the issue by consulting one of the following guides:

Plant Disease: Pathogens and Cycles
https://cropwatch.unl.edu/soybean-management/plant-disease

Wisconsin Field Crops Pathology
https://fyi.extension.wisc.edu/fieldcroppathology/soybean_pests_diseases/