

# What's lurking in the soil?

<b>Focus questions</b>	How do researchers determine the identity of organisms from pieces of DNA?
<b>Vocabulary</b>	Bioinformatics, barcodes

**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.

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.

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.

## 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 decides to use a soil and residue test from a diagnostic laboratory 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. 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  
[cropwatch.unl.edu/soybean-management/plant-disease](http://cropwatch.unl.edu/soybean-management/plant-disease)
- Wisconsin Field Crops Pathology  
[fyi.extension.wisc.edu/fieldcroppathology/soybean\\_pests\\_diseases/](http://fyi.extension.wisc.edu/fieldcroppathology/soybean_pests_diseases/)

# Sequence 1

## Procedure

1. Search the NCBI database to identify the following sequence of DNA.
2. Visit [ncbi.nlm.nih.gov](http://ncbi.nlm.nih.gov)
3. Click on 'BLAST' in the right column.
4. Click on 'Nucleotide blast'; paste in DNA sequence in the rectangle; submit; wait for result.  
*You can find sequences at [nourishthefuture.org/go/sequences](http://nourishthefuture.org/go/sequences).*

>Seq1

```
AGTCGGCGGGCTGGCTGCTGTGTGGCGGGCTCTATCATGGCGATTGGTTTGGGTCCTCCTCGTG
GGGAACTGGATCATGAGCCCACTTTTTAAACCCATTCTTAAATACTGAATATACTGTGGGGAC
GAAAGTCTCTGCTTTTAACTAGATAGCAACTTTTCAGCAGTGGATGTCTAGGCTCGCACATCGA
TGAAGAACGCTGCGAACTGCGATACGTAATGCGAATTGCAGGATTCAGTGAGTCATCGAAATT
TTGAACGCATATTGCACTTCCGGGTTAGTCCTGGGAGTATGCCTGTATCAGTGTCCGTACATC
AAACTTGGCTCTCTTCCCTTCCGTGTAGTCGGTGGATGGAGACGCCAGACGTGAGGTGTCTTGC
GGCGTGGCCTTCGGGCTGCCTGCGAGTCCCTTGAAATGTACTGAACTGTACTTCTCTTTGCTC
GAAAAGCGTGACGTTGTTGGTTGTGGAGGCTGCCTGTATGGCCAGTCGGCGACCGGTTTGTCT
GCTGCGGCGTTTAAATGGAGGAGTGTTTCGATTGCGGGTATGGTTGGCTTCGGCTGAACAATGCG
CTTATTGGATGCTTTTCCTGCTGTGGCGGTATGGGCTGGTGAACCGTAGCTGTGTGAGGCTTG
GCTTTTGAACCGGCGGTGTTGTTGC
```

## Reflection

1. Which organism does this sequence identify?
2. What is its classification (bacteria, fungus, nematode, plant, animal)?
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?

## Sequence 2

### Procedure

1. Search the NCBI database to identify the following sequence of DNA.
2. Visit [ncbi.nlm.nih.gov](https://ncbi.nlm.nih.gov)
3. Click on 'BLAST' in the right column.
4. Click on 'Nucleotide blast'; paste in DNA sequence in the rectangle; submit; wait for result.

>Seq2

```
AGTGATTCTATTCCCACCTACCTGCTGTCCTGTTGGGCTAGCGTTGGCACCACCAAATGCCCC
CGTCCGCTGATGGGCACAGGTCGTTGAGATGACTTGTGGACGCTGCCAACATTACGGGGCA
GCTGCCTCACGAGCCATGCTTTTGGGGTGCTTCCATACGTTGGAGCTGTGGTATACCGCTCAG
TGCTGCACATGTGAAAGCCTGTGTATGGCTGCTGCGTGGCAATGTGTGCGGTGGCGGGCCGCT
CGCTTGGCTGGTTCGCTGCGCCAATGTGGGATGCACGCTCGTGGGGCGACCTAACGGCTGTGC
TGGCGTCTGTGCGTCGTTGAGCGGTTGTTGTGGCAGGCACATAACACACTGACTGGGGATGGT
GGTTTCGTTCCCGGTCTTACGTGCCGTAAGTACGGGTGTGTTTGTGCTTGCTGCTACGTCCGT
GGCCGTGATGAGACGACGCGGTAGGGCCCGTGCTTGGCCTAGCACGTGGCTTAAGACTCAATG
AGTGTCAGCTCGGGCACCGCCAGCTTTTTCTTTTTTTTTCATTATTTTTTTTACACTTCTGT
TGAAGAATGAATTCTAGTCTTATCGGTGGATCACTCGGCTCGTGGATCGATGAAGAACGCAGC
CAACTGCGATAATTAGTGCGAACTGCAGAAACCTTGAACACAAAACATTGCAATGCACATTGC
GCCATTGGAGTTACATCCATTGGCACGCCTGGTTTCAAGGTCGTTACCATAAAAATGCACTGCTT
GTGCGTTGCTTCGTGGGATCATGTACTTGTACGTGTTCTTACGTTACTTGCTCAGCTCGGCTG
TGGGGTTTTGGTGTGCTGGCGCGAACTTGTGGTTCTAATTCGCGTTTTACGGACCGTAACTCG
GGCGCACCAATGCTTTGCATGCTGTGGCGGAGTGCCTGGATTACTGGCATTCTGCTTGATT
```

### 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?

## Sequence 3

### Procedure

1. Search the NCBI database to identify the following sequence of DNA.
2. Visit [ncbi.nlm.nih.gov](https://ncbi.nlm.nih.gov)
3. Click on 'BLAST' in the right column.
4. Click on 'Nucleotide blast'; paste in DNA sequence in the rectangle; submit; wait for result.

>Seq3

```
TGCGGAAGGATCATTGTTCGATGCCTCACAATCAGATTGACCCGCGAACTTGTTTATTCATCTA
CCGTTCGGGAGGGAGGGGATGACCACGGCGCCCCGTGCGCCCGGCCTCCTCGTCCTCGCGACAA
ACACAAACCCCGGCGCTTCGTGCGCCAAGGAACTCAAATCTGTTAAGTGCGACTCCCGGGGGC
CCGGAGACGGTGTCCCGCGGGAGTCGTACGACACAACATTTACATACAATGACTCTCGGCAA
CGGATATCTCGGCTCTTGCATCGATGAAGAACGTAGCGAAATGCGATACTTGGTGTGAATTGC
AGAATCCCGTGAACCATCGAGTCTTTGAACGCAAGTTGCGCCCGAAGCCATTAGGCCGAGGGC
ACGCCTGCCTGGGTGTACACATCGTTTCCCCAACGCAAACATGTAACAATGTTGCTGCGCGG
GGTGTATGCTGACCTCCCGCGAGCACCCGCCTCGTGGTTGGTTGAAATCTGGGTTTCATGGCCG
ACTTCGCCGTGATAAAATGGTGGATGAGCCACGCTCGAGACCAATCACGTGCGAGCCGGTCAG
TTCTGGACCCATCGACGACCCTTTGCGTGCACGCACGCTCCCAACGAGACCTCAGGTCAGGCG
GGGCTACCCGCTGAGTTTAAGCATATCAATAA
```

### 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?

## Sequence 4

### Procedure

1. Search the NCBI database to identify the following sequence of DNA.
2. Visit [ncbi.nlm.nih.gov](https://ncbi.nlm.nih.gov)
3. Click on 'BLAST' in the right column.
4. Click on 'Nucleotide blast'; paste in DNA sequence in the rectangle; submit; wait for result.

>Seq4

```
TGAATTAGGCCAACCCGGAACCTCTGCTCGGAGACGACCAAATCTACAACGTAGTTGTAACCGC
ACACGCATTTGTAATAATCTTCTTCATAGTAATACCAATCATAATTGGAGGATTTCGGTAACTG
ACTTGTTCCCTAATAATTGGTGCTCCCGATATAGCATTTCGCCGAATAAATAATATAAGCTT
CTGACTCCTCCCTCCCTCATTCTACTCTCCTCGCATCCTCTATAGTTGAAGCTGGGGCAGG
AACAGGCTGAACCGTGTACCCTCCCTTAGCAGGCAACCTAGCCCATGCAGGAGCTTCAGTAGA
TCTAACCATTTTCTCTTTACACTTAGCAGGAGTTTCCTCAATTTTAGGAGCCATCAACTTCAT
TACAACAATTATCAACATAAAGCCCCCGCAATGTCACAATACCAAACCCCTCTGTTTCGTATG
ATCCGTAATAATTACCGCCGTACTACTACTCTCGCTCCCTGTATTAGCAGCCGGGCATCAC
AATGCTATTAACAGACCGGAACCTAAATACA
```

### 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?

## Sequence 5

### Procedure

1. Search the NCBI database to identify the following sequence of DNA.
2. Visit [ncbi.nlm.nih.gov](https://ncbi.nlm.nih.gov)
3. Click on 'BLAST' in the right column.
4. Click on 'Nucleotide blast'; paste in DNA sequence in the rectangle; submit; wait for result.

>Seq5

```
CCTATATTTACTATTTGGTGCTTGAGCAGGTATAGTAGGAAGTGCCTTAAGCCTACTAATCCG
TGCTGAACTGGGTCAACCTGGGACTCTACTCGGAGATGATCAAATTTATAACGTAATTGTTAC
CGCACATGCATTCGTAATAATTTTCTTTATAGTTATACCAATTATAATTGGAGGATTCGGCAA
TTGACTTGTTCCATTAATAATTGGTGCTCCAGATATAGCATTCCCCCGAATAAATAACATAAG
CTTTTGACTTCTCCCTCCCTCTTTTTTATTACTTCTAGCATCATCTATAGTTGAAGCCGGAGC
AGGGACAGGCTGAACTGTTTATCCCCCTCTAGCTGGCAATCTAGCTCACGCAGGAGCTTCAGT
AGACCTAACTATTTTTTCTCTACACTTGGCGGGTGTCTCCTCGATTTTAGGAGCTATTAACCTT
TATTACAACAATTATCAACATAAAACCCCTGCTATATCACAATATCAAACCTCCTTTATTTGT
ATGATCTGTATTAATTACTGCCGTACTGCTACTTCTCTCACTCCCTGTATTAGCAGCTGGAAT
TACAATACTATTAACAGACCGAAATTTAAACACAACCTTTTTTCGACCCAGCAGGAGGCGGAGA
CCCCATCCTATATCAACACCTG
```

### 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?

# Rubric for self-assessment

Skill	Yes	No	Unsure
I see the advantage of using the NCBI database to identify organisms.			
I am able to identify an organism by using the NCBI database.			