

Finding a Gene in DNA

- What is a String2 min
- Understanding Strings3 min
- Developing an Algorithm5 min
- Positions in Strings8 min
- Translating into Code11 min
- Java Math8 min
- Programming Exercise: Finding a Gene and Web Links10 min
- Practice Quiz:** Finding a Gene in DNA6 questions

Finding All Genes in DNA

Debugging Code

Using the StorageResource Class

Review

Raluca Gordân

But first, I'd like you to think about the word strings.

What is a String

Have a question? Discuss this lecture in the week forums.



Interactive Transcript

Search Transcript

English

0:03

Hi, I'm Dr. Raluca Gordan. I'm a professor in the Duke Center for Genomic and Computational Biology and in the Department of Biostatistics and Bioinformatics.

0:12

My work is strongly based on designing and using computational algorithms, programs, and tools, and I want to tell you just a little bit about that.

0:21

But first, I'd like you to think about the word strings.

0:25

What does an orchestra conductor think about when she hears the word strings? What does a sailor think about the word strings? What does a pianist think about the word strings?

0:39

Well, I'm a genome scientist, so I will tell you what I think when I hear the word strings, and that is genomic strings. The genome of an organism stores all the genetic information necessary to build and maintain that organism. This genetic information is stored as a long list, or string, over the four-letter alphabet, A, T, C, and G.

1:02

These four characters correspond to the four DNA bases, adenine, thymine, cytosine, and guanine.

1:17

The sheer size of the genome makes it difficult, if not impossible, to analyze by hand. The human genome, for example, contains 3 billion characters. That is a million times more than the characters shown here. Thus, finding any information in the genome requires computational approaches.

1:36

In addition, the genome is complex and contains different types of information. Computational approaches are needed to find information, including genes, as you can see here.

1:49

Finding genes requires more than simply looking for the tags or codons that identify the start and end of a gene.

1:58

In addition to genes, such as the ones shown here in red, we also need to look for regulatory elements, as shown here. We do this with computational tools and techniques. These regulatory elements are shown here as simple letters representing nucleotides. But it's important to remember that these are actually bound by proteins, called transcription factors, which help turn genes on and off. My research is focused on identifying such regulatory elements in the human genome using various computational approaches. You will do similar things in this course, and this will prepare you to become a computer scientist or a genome scientist, depending on what your preferences are.

Downloads

Lecture Video mp4

Subtitles (English) WebVTT

Transcript (English) txt

Would you like to [help us translate](#) the transcript and subtitles into additional languages?