Q&A and Counting k-mers Example

Q&A

Counting k-mers

Counting k-mers

A common operation in computational genomics is finding the number of k-mers in a DNA sequence. This is the number of all possible substrings of length k in a given sequence.

Counting k-mers

A common operation in computational genomics is finding the number of k-mers in a DNA sequence. This is the number of all possible substrings of length k in a given sequence. For example, the sequence

<div align="center">

AAGAGT

</div>

has the following 2-mers:

Counting k-mers

A common operation in computational genomics is finding the number of k-mers in a DNA sequence. This is the number of all possible substrings of length k in a given sequence. For example, the sequence

AAGAGT

has the following 2-mers:

AA

Counting k-mers

A common operation in computational genomics is finding the number of k-mers in a DNA sequence. This is the number of all possible substrings of length k in a given sequence. For example, the sequence

<p style="text-align:center">AAGAGT</p>

has the following 2-mers:

AA

AG

Counting k-mers

A common operation in computational genomics is finding the number of k-mers in a DNA sequence. This is the number of all possible substrings of length k in a given sequence. For example, the sequence

<p style="text-align:center">AAGAGT</p>

has the following 2-mers:

AA

AG

GA

Counting k-mers

A common operation in computational genomics is finding the number of k-mers in a DNA sequence. This is the number of all possible substrings of length k in a given sequence. For example, the sequence

<p align="center">AAGAGT</p>

has the following 2-mers:

AA

AG   (twice)

GA

Counting k-mers

A common operation in computational genomics is finding the number of k-mers in a DNA sequence. This is the number of all possible substrings of length k in a given sequence. For example, the sequence

<p style="text-align:center">AAGA**GT**</p>

has the following 2-mers:

AA

AG (twice)

GA

GT

Counting k-mers

A common operation in computational genomics is finding the number of k-mers in a DNA sequence. This is the number of all possible substrings of length k in a given sequence. For example, the sequence

AAGAGT

has the following 2-mers: And the following 3-mers:

AA

AG   (twice)

GA

GT

Counting k-mers

A common operation in computational genomics is finding the number of k-mers in a DNA sequence. This is the number of all possible substrings of length k in a given sequence. For example, the sequence

AAGAGT

has the following 2-mers:

AA

AG (twice)

GA

GT

And the following 3-mers:

AAG

Counting k-mers

A common operation in computational genomics is finding the number of k-mers in a DNA sequence. This is the number of all possible substrings of length k in a given sequence. For example, the sequence

<div align="center">

AAGAGT

</div>

has the following 2-mers:

AA

AG   (twice)

GA

GT

And the following 3-mers:

AAG

AGA

Counting k-mers

A common operation in computational genomics is finding the number of k-mers in a DNA sequence. This is the number of all possible substrings of length k in a given sequence. For example, the sequence

AAGAGT

has the following 2-mers: And the following 3-mers:

AA   AAG

AG   (twice)   AGA

GA   GAG

GT

Counting k-mers

A common operation in computational genomics is finding the number of k-mers in a DNA sequence. This is the number of all possible substrings of length k in a given sequence. For example, the sequence

AAGAGT

has the following 2-mers: And the following 3-mers:

AA   AAG

AG   (twice)   AGA

GA   GAG

GT   AGT