

The Role of Kmers in Genome Assembly

Aengus McGuinness*, Erika Pedersen†

2021-07-23

Abstract

Kmers are pieces of a deoxyribonucleic acid of any length and are used extensively in genome assembly.

1 Motivation

At the beginning of our fellowship with the Institute for Computing in Research, we knew we wanted to explore the intersection of biology and computing in the field of bionformatics. With the guidance of our mentor, Dr. Joann Mudge, we explored different fields and applications of computing and gained a better understanding of the need for computing and innovation in the bioinformatics field.

2 Hypothesis

3 Essential Terminology

k-mers = A length of a DNA or RNA sequence of length 'k' that is generated from a longer, (usually duplicated) read of DNA or RNA. These k-mers are then used to assemble a genome using De Bruijn graphs, create k-mer sketches, or generate graphs to see how often k-mers occur. De Bruijn =

4

5

Kmers Kmers are peices of DNA of any length

*aengus82520@gmail.com

†erikalaraine2@gmail.com