Building compacted de Bruijn graph from 100 human genomes

Ilia Minkin[1] and Paul Medvedev[1]

Department of Computer Science and Engineering, The Pennsylvania State University, USA

1 Introduction

A paragraph stub. Discuss the importance of de Bruijn graphs [1] in assembly [cite assembly applications] and comparative genomics [cite comparative genomics applications].

A paragraph stub. Tell about compressed graph and its advantages[cite paper where it first appeared]. State that it is desirable to avoid construction of an ordinary graph first.

A paragraph stub. Notice that for pan-genome all available methods are based on suffix arrays/tress – they need a lot of memory.

2 Preliminaries

3 The Basic Algorithm

4 Parallelization Scheme

5 Effects of Bloom Filter Size and Parameter Selection

6 Results

7 Discussion

References

1. Bruijn, d.N.: A combinatorial problem. Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen. Series A 49(7), 758 (1946)