

August 1, 2016

To whom it may concern,

We are submitting this manuscript entitled *Variable-Order de Bruijn Graphs* for publication as a short paper in IEEE Transactions on Computational Biology and Bioinformatics.

Part of this manuscript was previously published in Proceedings of the Data Compression Conference, IEEE, 2015. We have extended it by describing a new external memory construction algorithm, and implemented it to allow us to demonstrate the scalability of our approach on much larger data sets.

The previous version only provided data for experiments on a 1.5 GB *E. coli* data set, and a 6.9 GB Human chromosome 14 data set. In this version, we repeat the experiments for the original data sets with our new construction method, and add two new data sets: a 26.7 GB Human genome data set, and a 70.3 GB Parrot data set. The new data sets exceed the RAM of our test machine, and produce graphs which entirely fill the RAM of our test machine.

We have also reported data for larger increases in order, and reported the peak RAM and disk usage during construction.

Sincerely,

Alex Bowe