

# Expandable de Novo Genome Assembler for Short-Read Sequence Data

Nikolay Vyahhi    Sergey Nurk    Anton Bankevich    Max Alekseyev    Pavel Pevzner

Algorithmic Biology Lab, Academic University of the Russian Academy of Sciences, St. Petersburg, Russia

<http://bioinf.spbau.ru/en/>

## Abstract

De novo genome sequence assembly is the essential step to reveal genomic sequences of different species world-wide. Currently there exists various genome assemblers for short-read NGS data, such as Velvet, SOAPdenovo, ALLPATH, ABySS and others. We present new open-source de Bruijn graph-based assembler currently in development on C++, which uses novel algorithmic ideas such as context-free graph approach and also have agile and expandable software architecture. It requires affordable amount of memory and computations while giving high quality results. It provides solid basis for single-cell and mammalian assemblers in the near future.