

Name	Year of publication	Reference	Method	Explicit for metagenomics	Description	Source code	Date of last update	Date of last release	Docker container	Included in the benchmark
Unicycler	2017	<a href="https://doi.org/10.1371/journal.pcbi.1005595">https://doi.org/10.1371/journal.pcbi.1005595</a>	dBg (multiple k-mer values)	no	Assembly pipeline for bacterial genomes. It can assemble illumina-only read sets where it functions as a SPAdes-optimiser.	<a href="https://github.com/rnwick/Unicycler">https://github.com/rnwick/Unicycler</a>	20/07/2020	16/08/2019	cimendes/unicycler:0.4.8-1	yes
MetaSPAdes	2017	<a href="https://dx.doi.org/10.1101%2Fgr.213959.116">https://dx.doi.org/10.1101%2Fgr.213959.116</a>	dBg (multiple k-mer values)	yes	MetaSPAdes is a specific metagenomic pipeline built on top of SPAdes. It constructs the de Bruijn graph of all reads using SPAdes, transforms it into the assembly graph using various graph simplification procedures, and reconstructs paths in the assembly graph that correspond to long genomic fragments within a metagenome.	<a href="https://github.com/ablab/spades/">https://github.com/ablab/spades/</a>	20/06/2020	02/05/2020	cimendes/spades:3.14.1-1	yes
SPAdes	2012	<a href="https://dx.doi.org/10.1089%2Fcmb.2012.0021">https://dx.doi.org/10.1089%2Fcmb.2012.0021</a>	dBg (multiple k-mer values)	no	A tool aiming to resolve uneven coverage in single cell genome data through multiple k-mer sizes of De Bruijn graphs. It starts with the smallest k-mer size and adds hypothetical k-mers to connect graph.	<a href="https://github.com/ablab/spades/">https://github.com/ablab/spades/</a>	20/06/2020	02/05/2020	cimendes/spades:3.14.1-1	yes
Minia	2013	<a href="https://doi.org/10.1186/1748-7188-9-22">https://doi.org/10.1186/1748-7188-9-22</a>	dBg (single k-mer value)	no	The assembly is performed on a data structure based on unitigs produced by the BCALM software and using graph simplifications that are heavily inspired by the SPAdes assembler.	<a href="https://github.com/GATB/minia">https://github.com/GATB/minia</a>	07/06/2020	07/06/2020	cimendes/minia:3.3.4-1	yes
GATB-Minia Pipeline	unpublished	unpublished	dBg (multiple k-mer values)	yes	The pipeline consists of Bloccoo for error correction, Minia 3 for contigs assembly, which is based on the BCALM2 assembler, and the BEST for scaffolding. It was developed to extend Minia assembler to use multiple k-mer values.	<a href="https://github.com/GATB/gatb-minia-pipeline">https://github.com/GATB/gatb-minia-pipeline</a>	31/05/2020	--	cimendes/gatb-minia-pipeline:31.05.2020-1	yes
bcalm2	2016	<a href="https://doi.org/10.1093/bioinformatics/btw279">https://doi.org/10.1093/bioinformatics/btw279</a>	dBg (single k-mer value)	no	Fast and low memory algorithm for graph compaction, consisting of three stages: careful distribution of input k-mers into buckets, parallel compaction of the buckets, and a parallel reunification step to glue together the compacted strings into unitigs.	<a href="https://github.com/GATB/bcalm">https://github.com/GATB/bcalm</a>	22/05/2020	22/05/2020	cimendes/bcalm:2.2.3-1	yes
Skesa	2018	<a href="https://doi.org/10.1186/s13059-018-1540-z">https://doi.org/10.1186/s13059-018-1540-z</a>	dBg (multiple k-mer values)	no	Skesa was designed for assembling reads of microbial genomes sequenced using Illumina. The heuristics used are designed to reduce the effect of low-level contamination and strand specific errors in Illumina sequencing on the quality of the assembly.	<a href="https://github.com/ncbi/SKESA/releases">https://github.com/ncbi/SKESA/releases</a>	14/03/2020	10/03/2019	cimendes/skesa:2.4.0-1	yes
MEGAHIT	2015	<a href="https://doi.org/10.1093/bioinformatics/btv033">https://doi.org/10.1093/bioinformatics/btv033</a>	dBg (multiple k-mer values)	yes	MEGAHIT is a de novo assembler for assembling large and complex metagenomics data in a time- and cost-efficient manner. It makes use of succinct de Bruijn graph, with a multiple k-mer size strategy. In each iteration, MEGAHIT clears potentially erroneous edges by removing tips, merging bubbles and removing low local coverage edges, specially useful for metagenomics which suffers from non-uniform sequencing depths.	<a href="https://github.com/youtan/megahit">https://github.com/youtan/megahit</a>	15/10/2019	15/10/2019	cimendes/megahit-assembler:1.2.9-1	yes
PANDaseq	2012	<a href="https://doi.org/10.1186/1471-2105-13-31">https://doi.org/10.1186/1471-2105-13-31</a>	OLC	no	PANDASEQ is a program to align illumina reads, optionally with PCR primers embedded in the sequence, and reconstruct an overlapping sequence	<a href="https://github.com/neufeld/pandaseq">https://github.com/neufeld/pandaseq</a>	30/08/2018	03/03/2017	cimendes/pandaseq:2.11-1	yes
VelvetOptimizer	unpublished	unpublished	dBg (multiple k-mer values)	no	The VelvetOptimizer is designed to run as a wrapper script for the Velvet assembler and to assist with optimising the assembly. It searches a supplied hash value range for the optimum, estimates the expected coverage and then searches for the optimum coverage cutoff. It uses Velvet's internal mechanism for estimating insert lengths for paired end libraries. It can optimise the assemblies by either the default optimisation condition or by a user supplied one. It outputs the results to a subdirectory and records all its operations in a logfile.	<a href="https://github.com/seemann/VelvetOptimizer">https://github.com/seemann/VelvetOptimizer</a>	21/12/2017	03/08/2017	cimendes/velvetoptimiser:2.2.6-1	yes
IDBA-UD	2012	<a href="https://doi.org/10.1093/bioinformatics/bts174">https://doi.org/10.1093/bioinformatics/bts174</a>	dBg (multiple k-mer values)	yes	It's an assembler for assembling reads from single-cell sequencing or metagenomic sequencing technologies with uneven sequencing depths. It employs multiple depth relative thresholds to remove erroneous k-mers in both low-depth and high-depth regions. The technique of local assembly with paired-end information is used to solve the branch problem of low-depth short repeat regions. To speed up the process, an error correction step is conducted to correct reads of high-depth regions that can be aligned to high confidence contigs.	<a href="https://github.com/oneknighthp/idba">https://github.com/oneknighthp/idba</a>	31/12/2016	11/07/2016	cimendes/idba:31.12.2016-3	yes
Velour	2011	<a href="http://hdl.handle.net/2142/24291">http://hdl.handle.net/2142/24291</a>	dBg (single k-mer value)	no	Velour avoids the large memory requirement exhibited by other assemblers by intelligently but efficiently caching portions of the assembly process on disk. Despite this memory optimization, the solution quality does not suffer and remains comparable to other popular assemblers	<a href="https://github.com/jcook/velour">https://github.com/jcook/velour</a>	13/05/2016	--		no
Metassembler	2015	<a href="https://doi.org/10.1186/s13059-015-0764-4">https://doi.org/10.1186/s13059-015-0764-4</a>	OLC	yes	Metassembler combines multiple whole genome de novo assemblies into a combined consensus assembly using the best segments of the individual assemblies.	<a href="https://sourceforge.net/projects/metassembler/">https://sourceforge.net/projects/metassembler/</a>	29/09/2015	23/06/2015		no
SoapDeNovo2	2012	<a href="https://doi.org/10.1186/2047-217X-1-18">https://doi.org/10.1186/2047-217X-1-18</a>	dBg (single k-mer value)	no	SOA-Pdenovo2 is made up of six modules that handle read error correction, de Bruijn graph construction, contig assembly, paired-end reads mapping, scaffold construction, and gap closure.	<a href="https://sourceforge.net/projects/soapdenovo2/">https://sourceforge.net/projects/soapdenovo2/</a>	07/08/2015	17/03/2015		no
Ray	2010	<a href="https://dx.doi.org/10.1089%2Fcmb.2009.0238">https://dx.doi.org/10.1089%2Fcmb.2009.0238</a>	dBg (single k-mer value)	no	A de Bruijn graph de novo assembler that defines specific subsequences, called seeds, and for each of them, the algorithm extends it into a contig. Defined heuristics control the extension process in such a way that the process is stopped if, at some point, the family of reads does not clearly indicate the direction of the extension.	<a href="https://sourceforge.net/projects/denovoassembler/files/">https://sourceforge.net/projects/denovoassembler/files/</a>	12/02/2014	12/02/2014		no
Ray Meta	2012	<a href="https://doi.org/10.7717/peerj.196">https://doi.org/10.7717/peerj.196</a>	dBg (single k-mer value)	yes	An extension of the Ray assembler, it applies no graph partitioning nor use a single peak for k-mer coverage, mean, and peak coverage are specific for each seed path. It employs heuristics-based graph transversal and graph is coloured according to an expected taxonomic profile.	<a href="https://sourceforge.net/projects/denovoassembler/files/">https://sourceforge.net/projects/denovoassembler/files/</a>	12/02/2014	12/02/2014		no
PRICE	2013	<a href="https://dx.doi.org/10.1534%2Fg3.113.005967">https://dx.doi.org/10.1534%2Fg3.113.005967</a>	Hybrid (single k-mer value)	yes	Targeted assembly strategy involved iteratively repeated execution of three steps: (1) for each seed contig, identification of a subset of data that could be assembled to expand the seed contig; (2) performance of an assembly on each seeded dataset; and (3) elimination of redundancy from the products of individual assembly jobs through performance of a meta-assembly on those products	<a href="https://sourceforge.net/projects/pricedenovo/">https://sourceforge.net/projects/pricedenovo/</a>	29/05/2013	08/05/2013		no
BBAP	2017	<a href="https://dx.doi.org/10.1186%2Fs12859-017-1630-z">https://dx.doi.org/10.1186%2Fs12859-017-1630-z</a>	OLC	yes	Assembly pipeline for the assembly of highly polymorphic metagenomic NGS data sets through a unique BLAST-based greedy algorithm.	<a href="http://homepage.ntu.edu.tw/~youylin/BBAP.html">http://homepage.ntu.edu.tw/~youylin/BBAP.html</a>	NA	01/12/2015		no
MetaVelvet-SL	2015	<a href="https://dx.doi.org/10.1093%2Fdnare%2Fdsu041">https://dx.doi.org/10.1093%2Fdnare%2Fdsu041</a>	dBg (single k-mer value)	yes	This De Bruijn graph assembler is similar to MetaVelvet, assembling data by populating a structure estimated from the coverage of nodes (poisson distributions). The De Bruijn graph is partitioned into hypothetical subgraphs (possible different species) using these distributions as a guide. Differently from MetaVelvet, MetaVelvet-SL identifies chimeric contigs through a support-vector machine (SVM) trained on paired-end, coverage, and contig lengths for each dinucleotide and passed on to the De Bruijn graph for decomposition.	<a href="http://metavelvet.dna.bio.keio.ac.jp/MSL.html">http://metavelvet.dna.bio.keio.ac.jp/MSL.html</a>	NA	01/01/2015		no
Velvet	2008	<a href="https://dx.doi.org/10.1101%2Fgr.074492.107">https://dx.doi.org/10.1101%2Fgr.074492.107</a>	dBg (single k-mer value)	no	Velvet manipulates these de Bruijn graphs efficiently to both eliminate errors and resolve repeats. These two tasks are done separately: first, the error correction algorithm merges sequences that belong together, then the repeat solver separates paths sharing local overlaps.	<a href="https://www.ebi.ac.uk/~zerbino/velvet/">https://www.ebi.ac.uk/~zerbino/velvet/</a>	NA	15/08/2014		no
xGenovo	2013	<a href="https://peerj.com/articles/196/">https://peerj.com/articles/196/</a>	OLC	yes	Extended Genovo by adding a bonus parameter in the Chinese Restaurant Process used to get prior accounts for the unknown number of genomes in the sample, intending for a pair of reads to be in the same contig as an effort to solve chimera contig case. Use of relative distance for the number of trials in the symmetric geometric distribution instead of using distance between the offset and the center of contig used in Genovo.	<a href="http://xgenovo.dna.bio.keio.ac.jp/download">http://xgenovo.dna.bio.keio.ac.jp/download</a>	NA	01/01/2013		no
MetaVelvet	2012	<a href="https://dx.doi.org/10.1093%2Fnar%2Fgks678">https://dx.doi.org/10.1093%2Fnar%2Fgks678</a>	dBg (single k-mer value)	yes	A modified and extended version of the single-genome and de Bruijn-graph based assembler, Velvet, for de novo metagenomic assembly. A decomposing de Bruijn graph constructed from mixed short reads into individual sub-graphs. Building scaffolds is based on every decomposed de Bruijn sub-graph as isolated species genome.	<a href="http://metavelvet.dna.bio.keio.ac.jp/src/">http://metavelvet.dna.bio.keio.ac.jp/src/</a>	NA	29/05/2012		no
Meraculous	2011	<a href="https://doi.org/10.1371/journal.pone.0023501">https://doi.org/10.1371/journal.pone.0023501</a>	dBg (single k-mer value)	no	Meraculous relies on an efficient and conservative traversal of the subgraph of the k-mer (deBruijn) graph of oligonucleotides with unique high quality extensions in the dataset, avoiding an explicit error correction step as used in other short-read assemblers.	<a href="ftp://ftp.jp-psf.org/pub/UGL_data/meraculous/">ftp://ftp.jp-psf.org/pub/UGL_data/meraculous/</a>	NA	20/09/2011		no
Genovo	2010	<a href="https://doi.org/10.1007/978-3-642-12683-3_22">https://doi.org/10.1007/978-3-642-12683-3_22</a>	OLC	yes	Genovo implements a Chinese restaurant process prior that accounts for the unknown number of genomes in the sample. Inference is made by applying a series of hill-climbing steps iteratively until convergence.	<a href="https://cs.stanford.edu/genovo.">https://cs.stanford.edu/genovo.</a>	NA	--		no