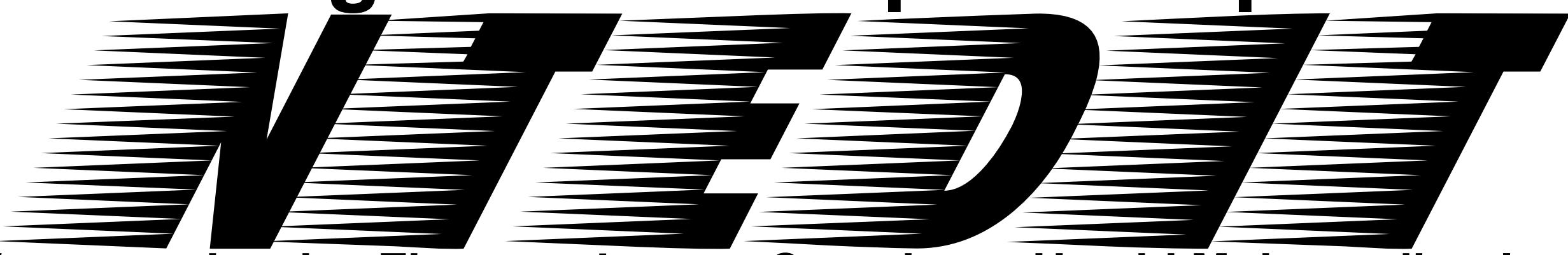
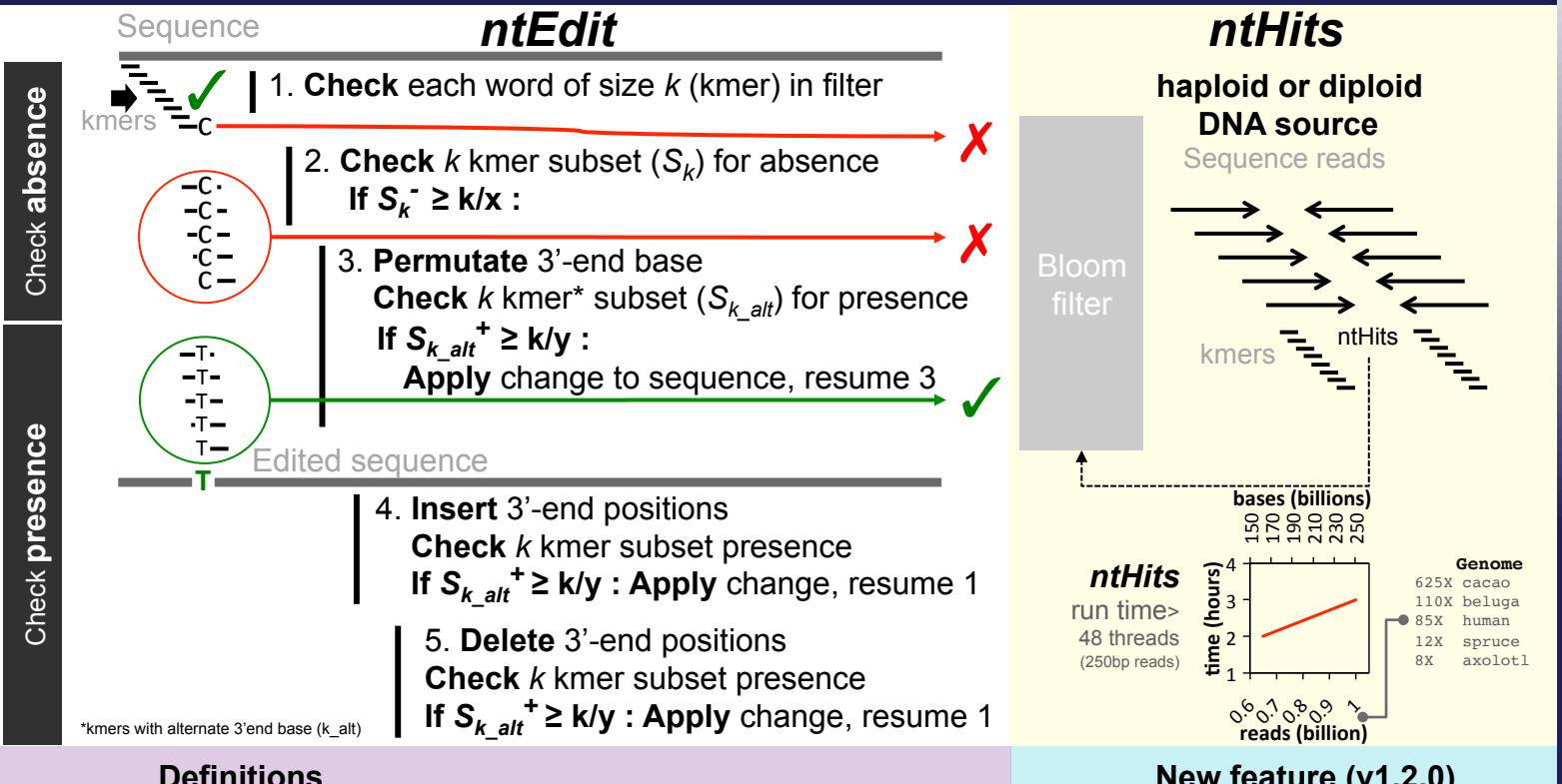


scalable genome sequence polishing

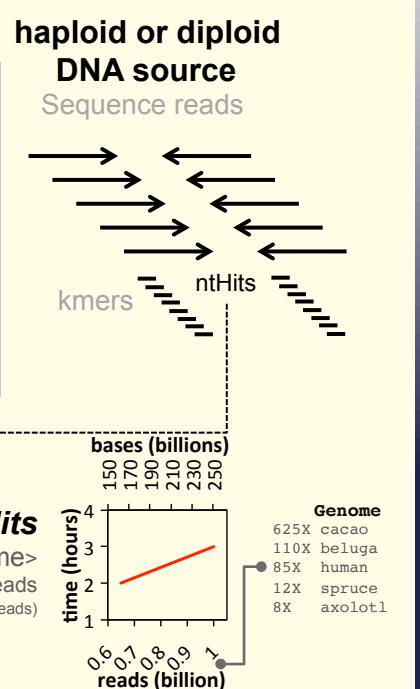


René Warren • Jessica Zhang • Lauren Coombe • Hamid Mohamadi • Inanç Birol

Method

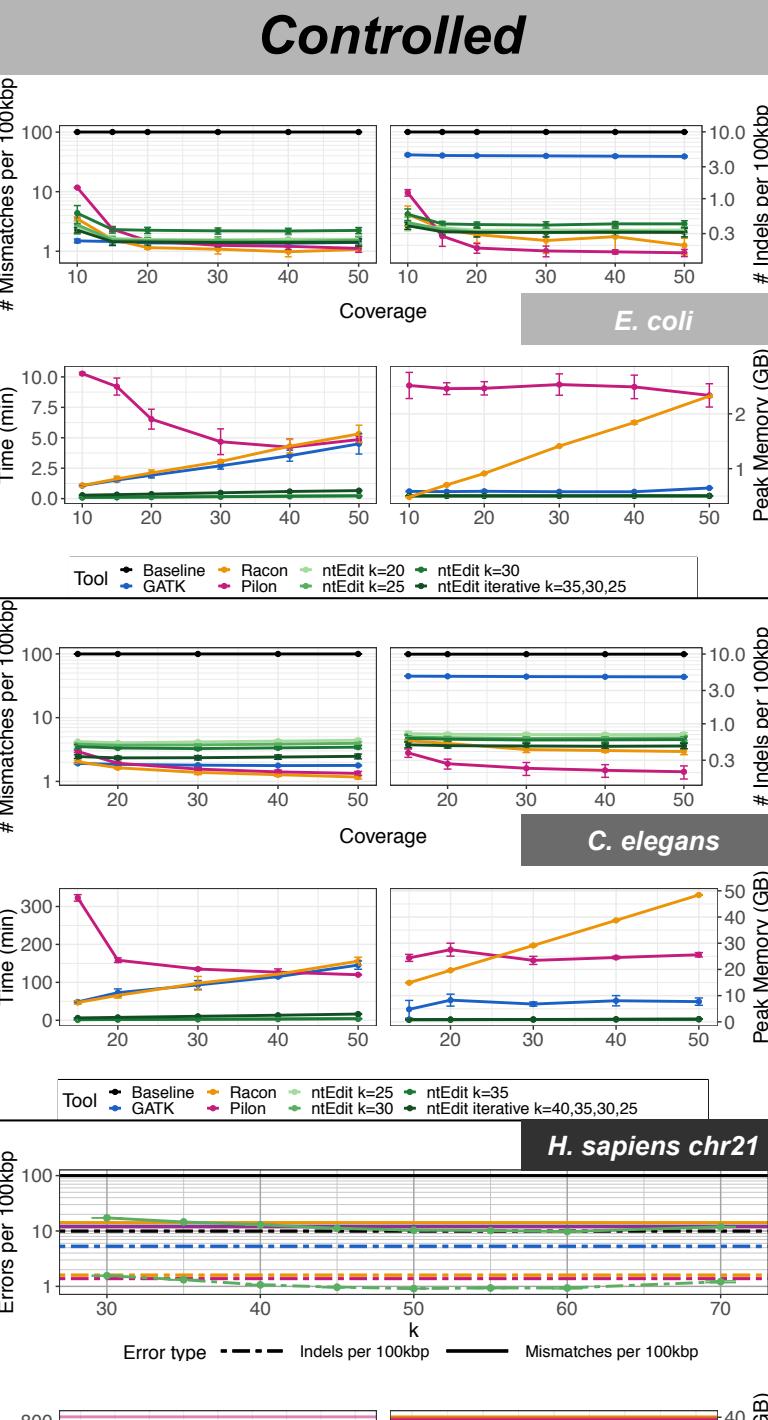


ntHits



Results

Controlled



Experimental

I. Polishing

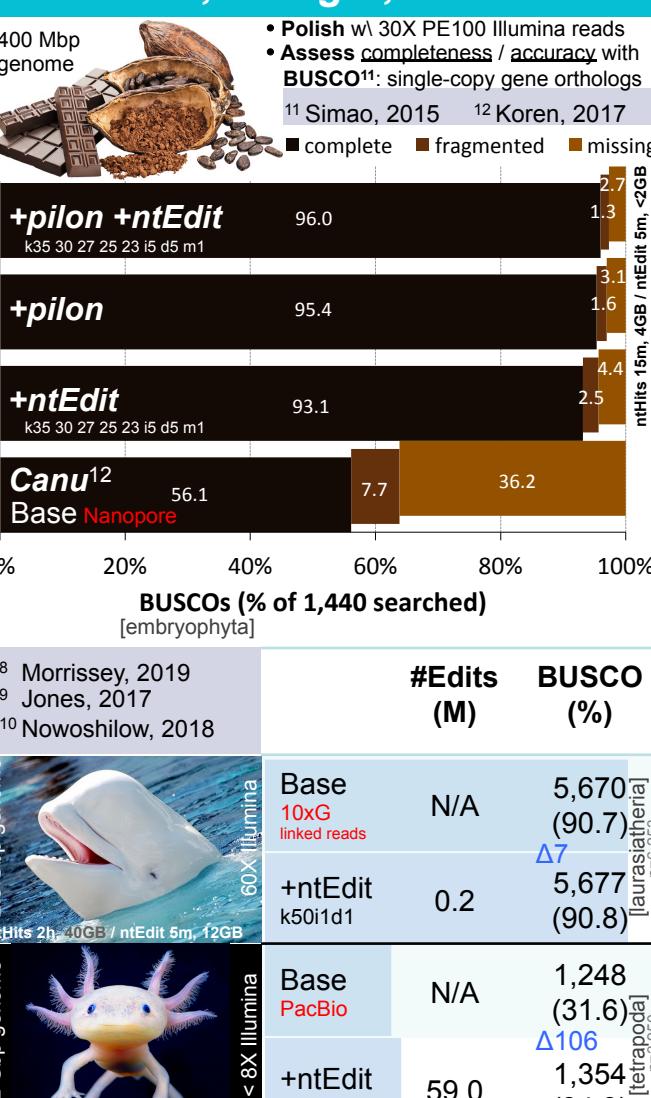
Human*

Polish w/ 54X Illumina	Time**	Edits (M)	BUSCO (%)	Baseline
GATK ⁴	41h45	0.97	5,654 (91.3)	7
ntEdit***	2h10	0.95	5,670 (91.6)	23
Racon ⁵	45h54	N/A	5,681 (91.7)	34
GATK	42h21	2.66	5,285 (85.4)	0
ntEdit***	2h10	3.63	5,651 (91.3)	366
Racon	40h55	N/A	5,670 (91.6)	5,285 (85.4)

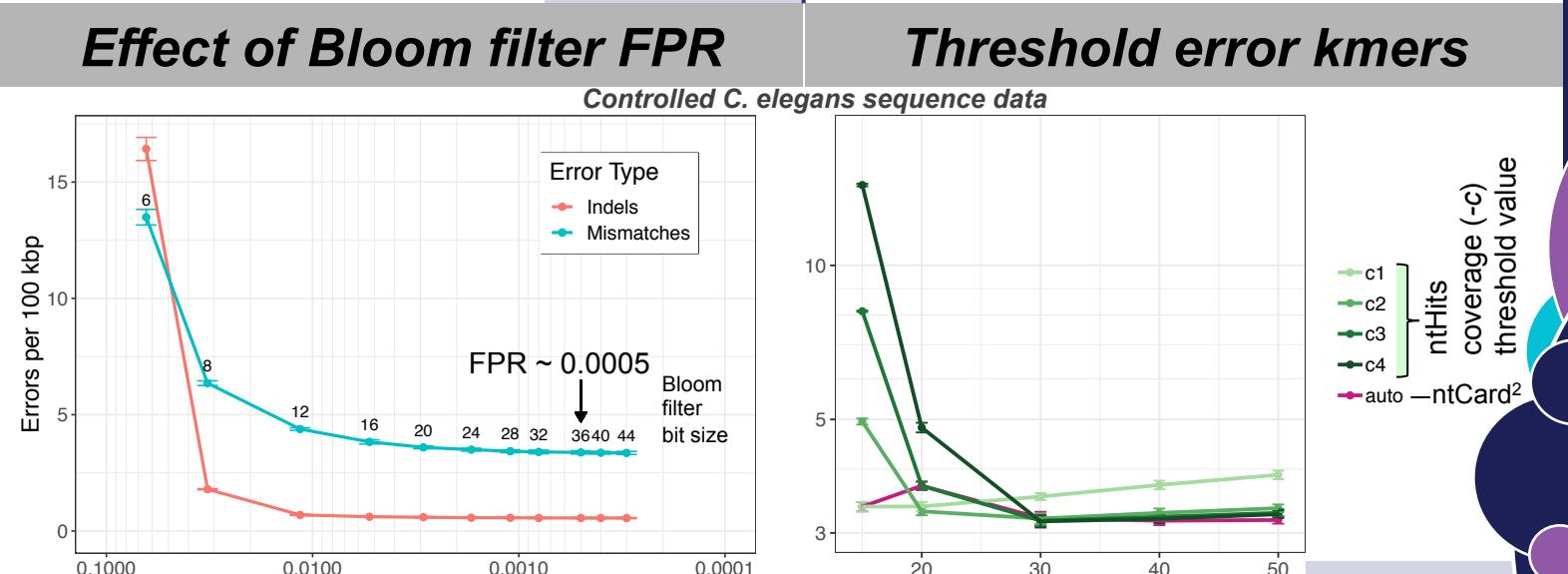
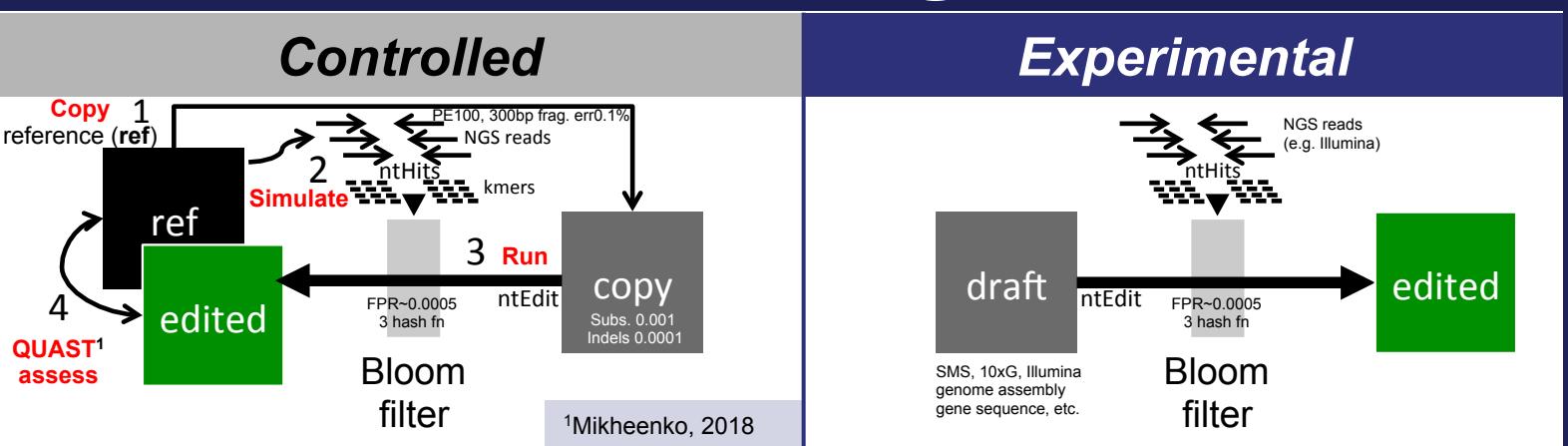
⁴McKenna, 2010 ⁵Vaser, 2017 ⁶Jain, 2018 ⁷Pendleton, 2015

*Single Molecule Sequencing draft genomes **Time for pipeline ***15GB RAM ntHits 40GB

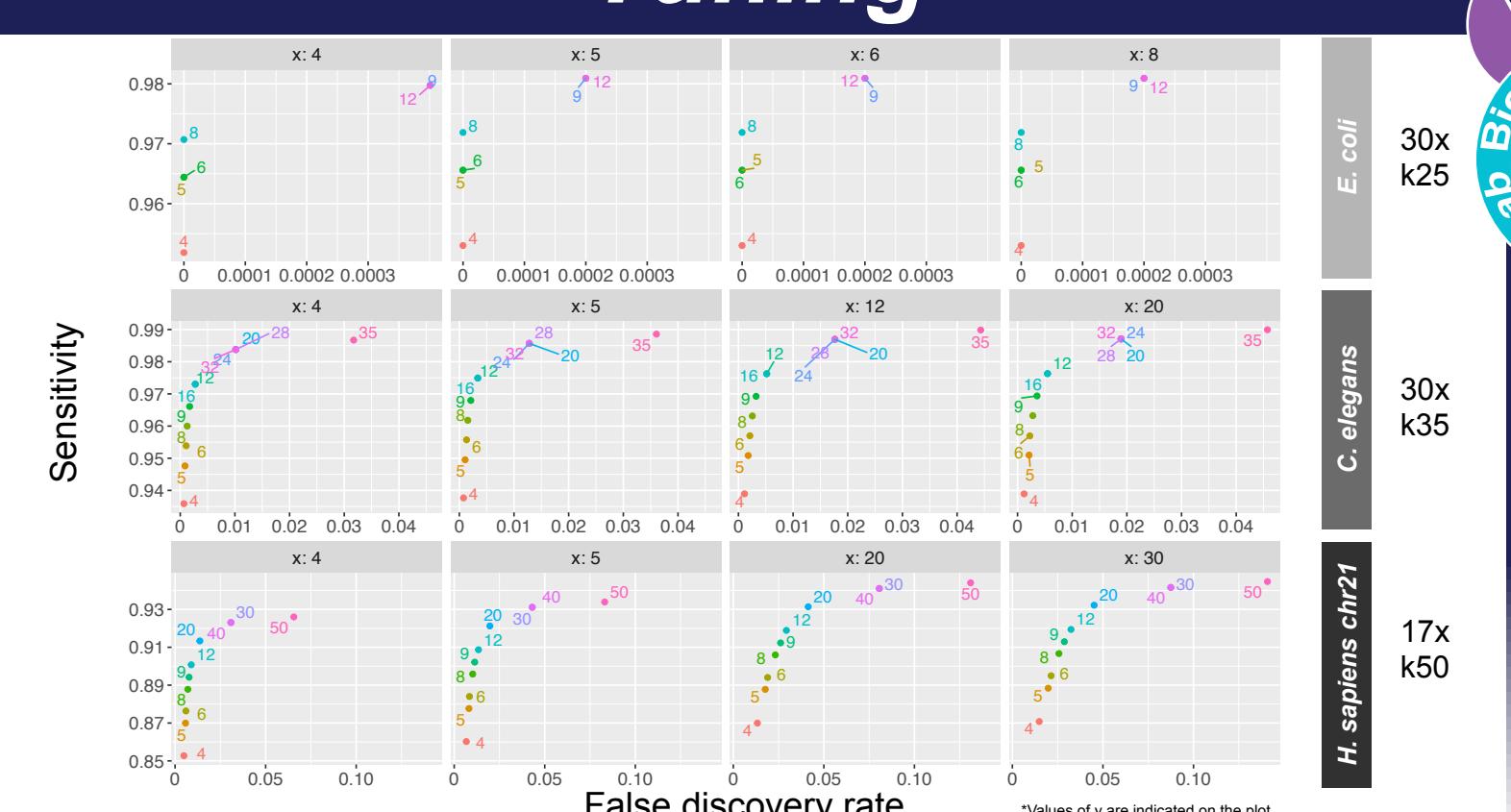
Cacao⁸, Beluga⁹, Axolotl¹⁰



Testing

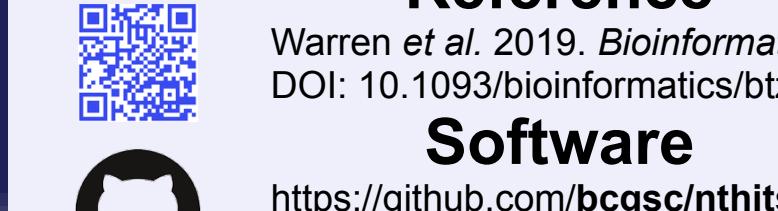
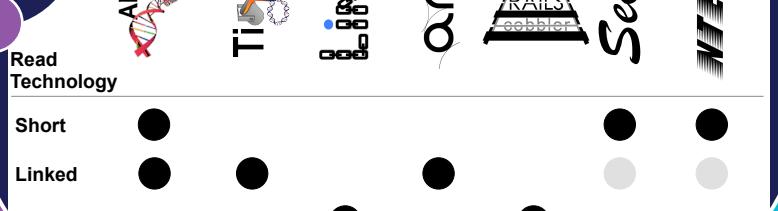
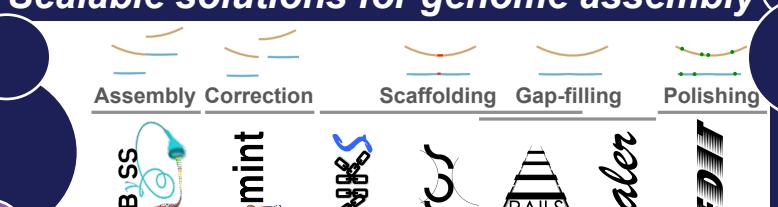


Tuning



Summary

Scalable solutions for genome assembly



Acknowledgements

Reference

Warren et al. 2019. *Bioinformatics*. DOI: 10.1093/bioinformatics/btz400

Software

<https://github.com/bcgsc/nthits>
<https://github.com/bcgsc/ntedit>

Funding

NIH

GenomeBritishColumbia

GenomeCanada

CANADA'S MICHAEL SMITH GENOME SCIENCES CENTRE

<http://birol-lab.ca>

<http://renewarren.ca>

BC CANCER FOUNDATION