

De novo genome assembly

Computational Genomics | Lecture 14

Tom Harrop

melbournebioinformatics.org.au

Research article | [Open Access](#) | Published: 08 April 2019

A hybrid de novo genome assembly of the honeybee, *Apis mellifera*, with chromosome-length scaffolds

[Andreas Wallberg](#), [Ignas Bunikis](#), [Olga Vinnere Petterson](#), [Mai-Britt Mosbech](#), [Anna K. Childers](#), [Jay D. Evans](#), [Alexander S. Mikheyev](#), [Hugh M. Robertson](#), [Gene E. Robinson](#) & [Matthew T. Webster](#) 

[BMC Genomics](#) 20, Article number: 275 (2019) | [Cite this article](#)

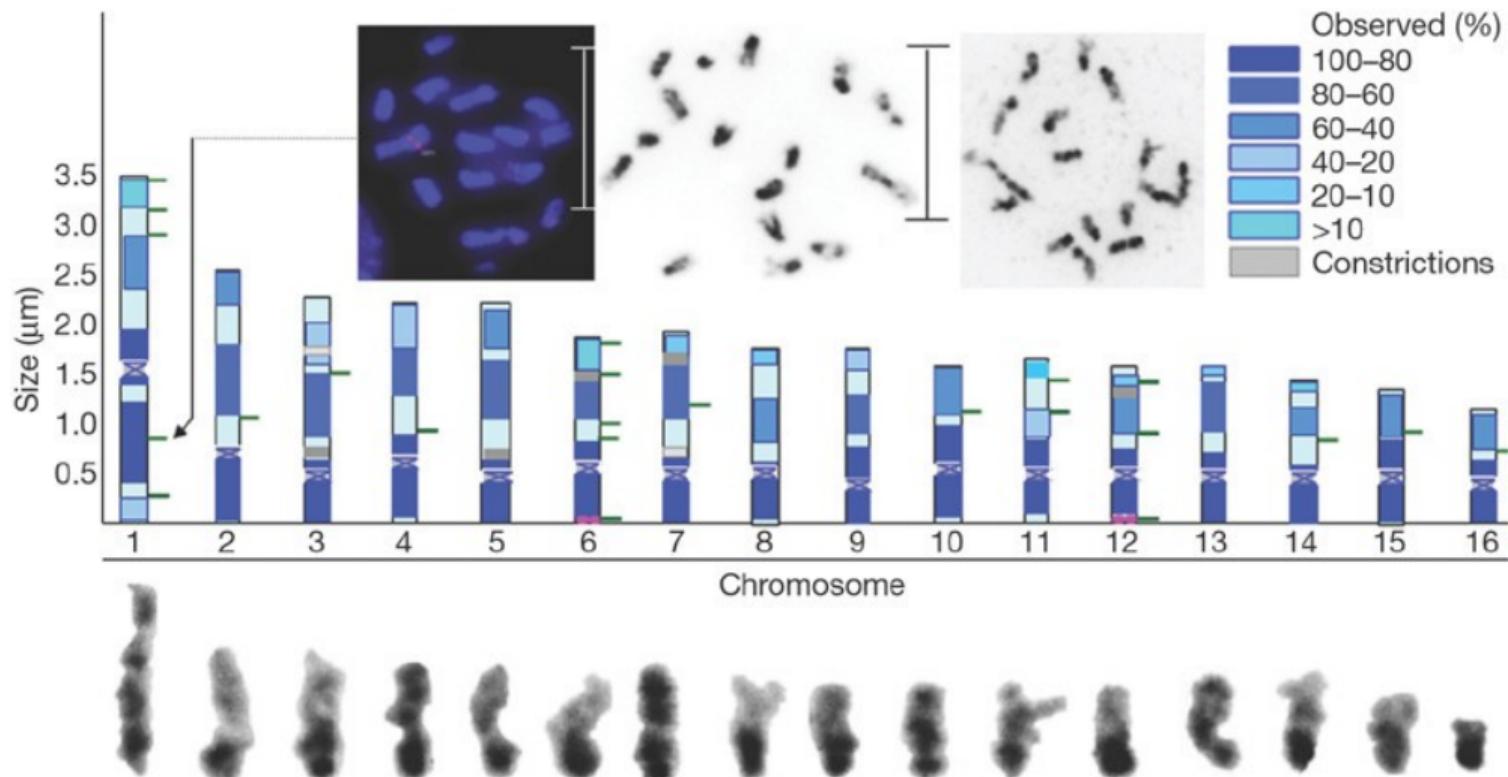
7651 Accesses | 32 Citations | 26 Altmetric | [Metrics](#)

Apis mellifera: Western honey bee



Ivar Leidus via Wikimedia Commons

First genome assembly: Amel_4.0



A deficit of detoxification enzymes: pesticide sensitivity and environmental response in the honeybee

C. Cladanos, H. Ranson, R. M. Johnson, S. Biswas, M. A. Schuler, M. R. Berenbaum, R. Feyereisen, J. G. Oakeshott

First published: 27 October 2006 | <https://doi.org/10.1111/j.1365-2583.2006.00672.x> |
Citations: 416

✉ C. Cladanos, Research School of Biological Sciences, Australian National University, Canberra, ACT 0200, Australia. Fax: +61(2)61253808; e-mail: charles.cladanos@anu.edu.au

Re-use of this article is permitted in accordance with the Creative Commons Deed, Attribution 2.5, which does not permit commercial exploitation.

≡ SECTIONS

 PDF  TOOLS  SHARE

Abstract

The honeybee genome has substantially fewer protein coding genes ($\approx 11\,000$ genes) than *Drosophila melanogaster* ($\approx 13\,500$) and *Anopheles gambiae* ($\approx 14\,000$). Some of the most marked differences occur in three superfamilies

First major upgrade: Amel_4.5

Research article | [Open Access](#) | Published: 30 January 2014

Finding the missing honey bee genes: lessons learned from a genome upgrade

[Christine G Elsik](#) , [Kim C Worley](#) , [Anna K Bennett](#), [Martin Beye](#), [Francisco Camara](#), [Christopher P Childers](#), [Dirk C de Graaf](#), [Griet Debysen](#), [Jixin Deng](#), [Bart Devreese](#), [Eran Elhaik](#), [Jay D Evans](#), [Leonard J Foster](#), [Dan Graur](#), [Roderic Guigo](#), [HGSC production teams](#), [Katharina Jasmin Hoff](#), [Michael E Holder](#), [Matthew E Hudson](#), [Greg J Hunt](#), [Huaiyang Jiang](#), [Vandita Joshi](#), [Radhika S Khetani](#), [Peter Kosarev](#), [Christie L Kovar](#), [Jian Ma](#), [Ryszard Maleszka](#), [Robin F A Moritz](#), [Monica C Munoz-Torres](#), [Terence D Murphy](#), [Donna M Muzny](#), [Irene F Newsham](#), [Justin T Reese](#), [Hugh M Robertson](#), [Gene E Robinson](#), [Olav Rueppell](#), [Victor Solovyev](#), [Mario Stanke](#), [Eckart Stolle](#), [Jennifer M Tsuruda](#), [Matthias Van Vaerenbergh](#), [Robert M Waterhouse](#), [Daniel B Weaver](#), [Charles W Whitfield](#), [Yuanqing Wu](#), [Evgeny M Zdobnov](#), [Lan Zhang](#), [Dianhui Zhu](#) & [Richard A Gibbs](#) [-Show fewer authors](#)

[BMC Genomics](#) 15, Article number: 86 (2014) | [Cite this article](#)

15k Accesses | 224 Citations | 33 Altmetric | [Metrics](#)

Hybrid approach: Amel_HAv3

Research article | [Open Access](#) | Published: 08 April 2019

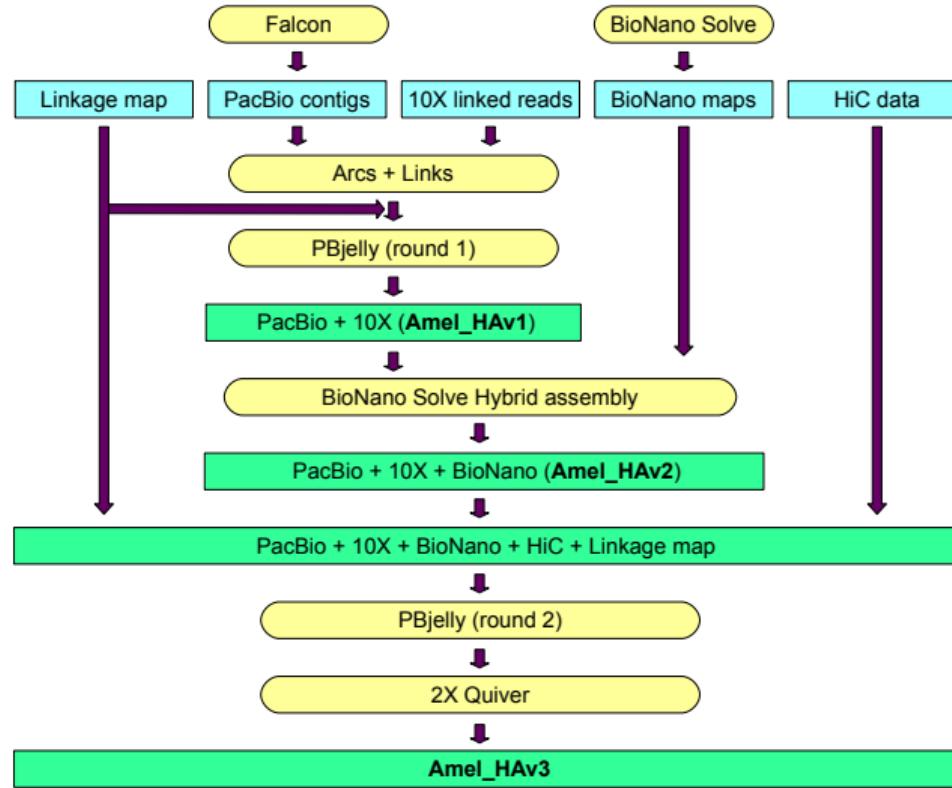
A hybrid de novo genome assembly of the honeybee, *Apis mellifera*, with chromosome-length scaffolds

[Andreas Wallberg](#), [Ignas Bunikis](#), [Olga Vinnere Pettersson](#), [Mai-Britt Mosbech](#), [Anna K. Childers](#), [Jay D. Evans](#), [Alexander S. Mikheyev](#), [Hugh M. Robertson](#), [Gene E. Robinson](#) & [Matthew T. Webster](#) 

[BMC Genomics](#) **20**, Article number: 275 (2019) | [Cite this article](#)

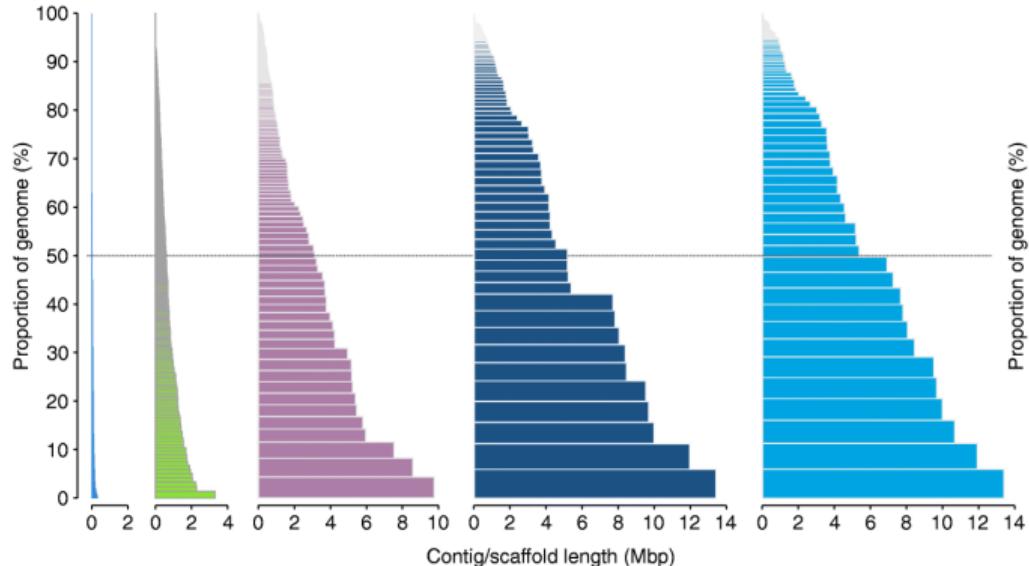
7651 Accesses | **32** Citations | **26** Altmetric | [Metrics](#)

Kitchen sink genome assembly

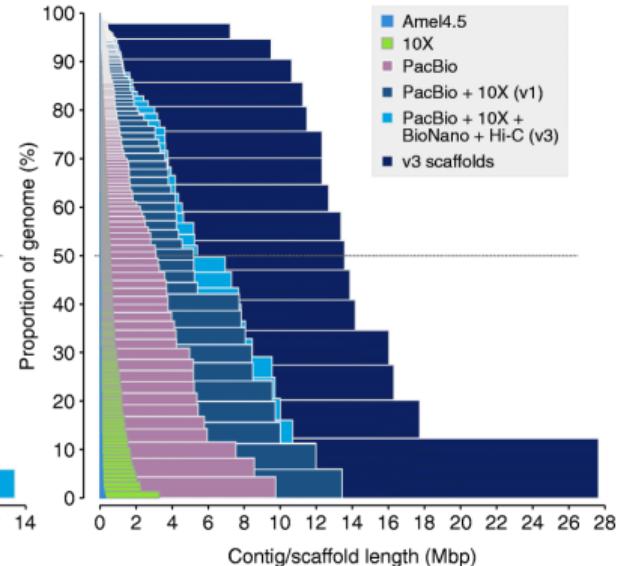


Contiguity is much better

A

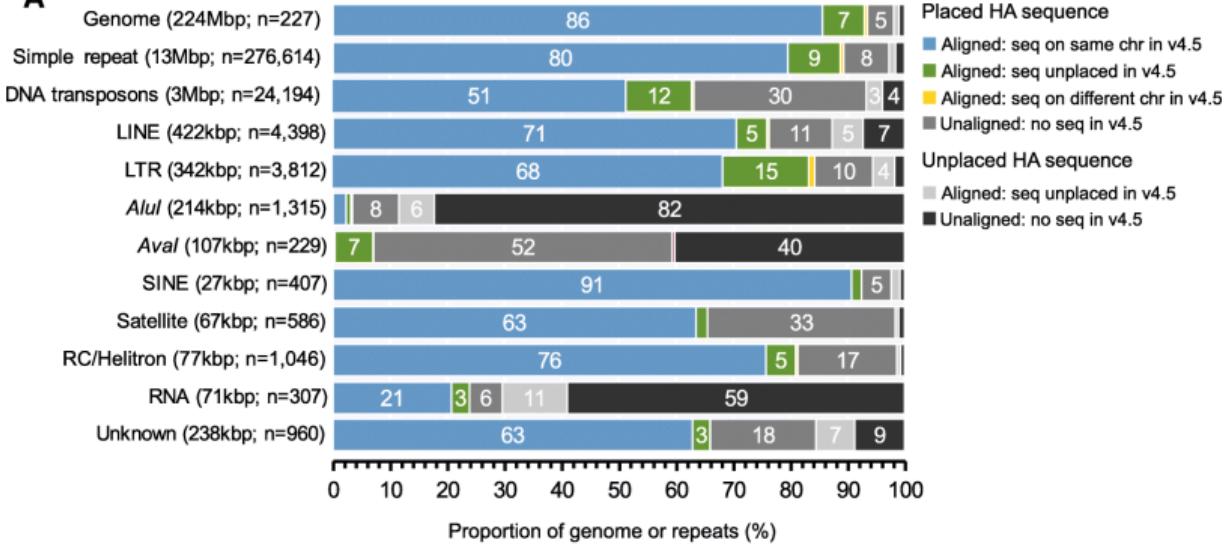


B

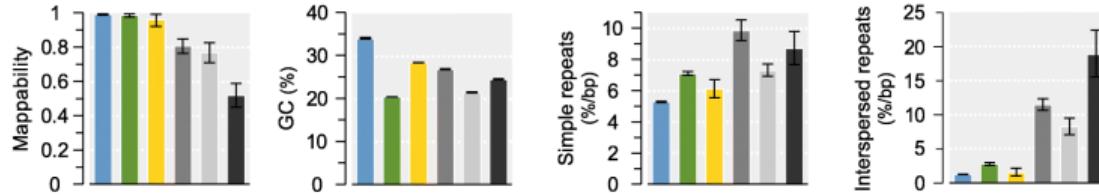


10% increase in genome size, 21% increase in repeat size

A



B

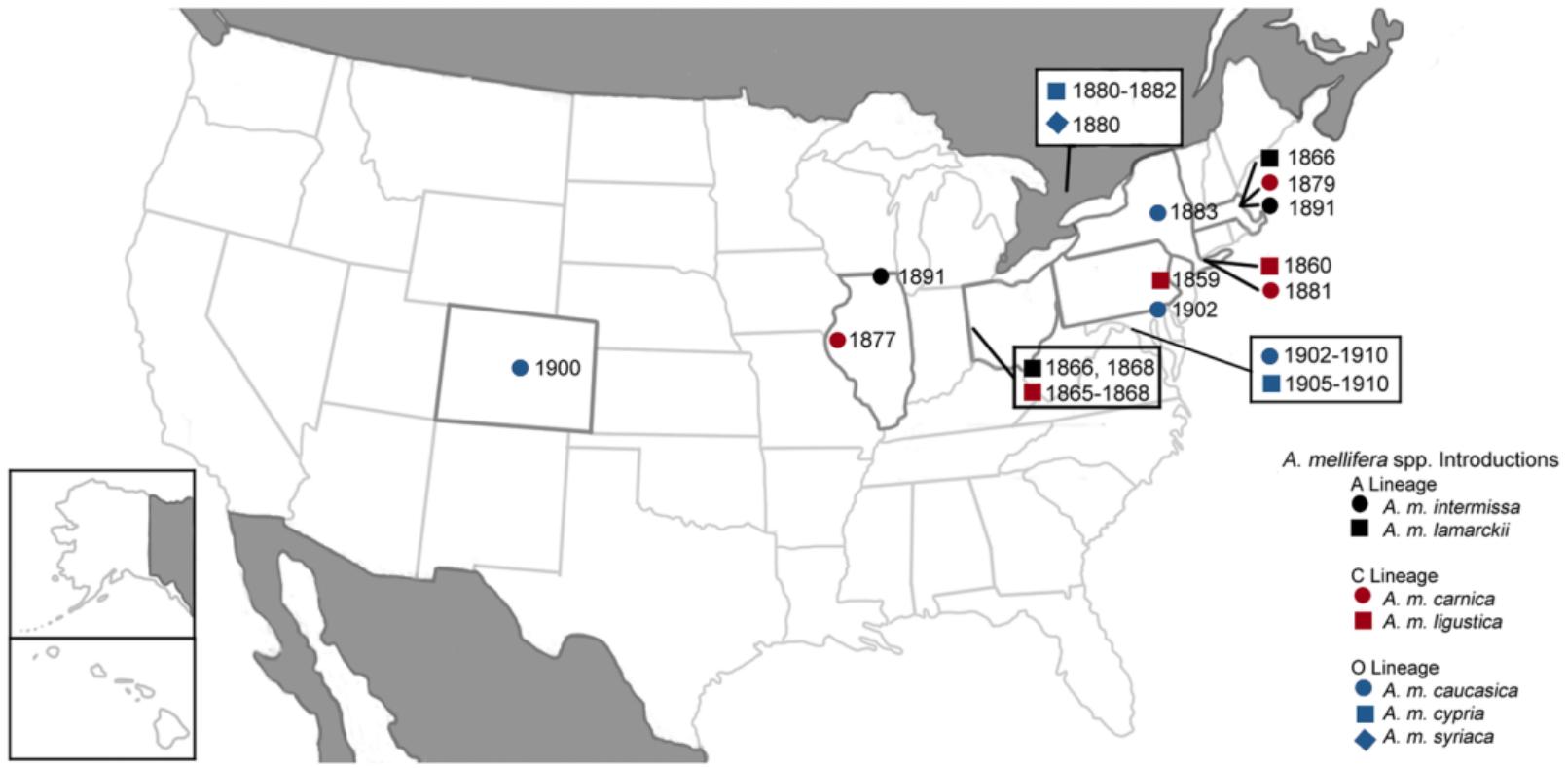


Some metrics

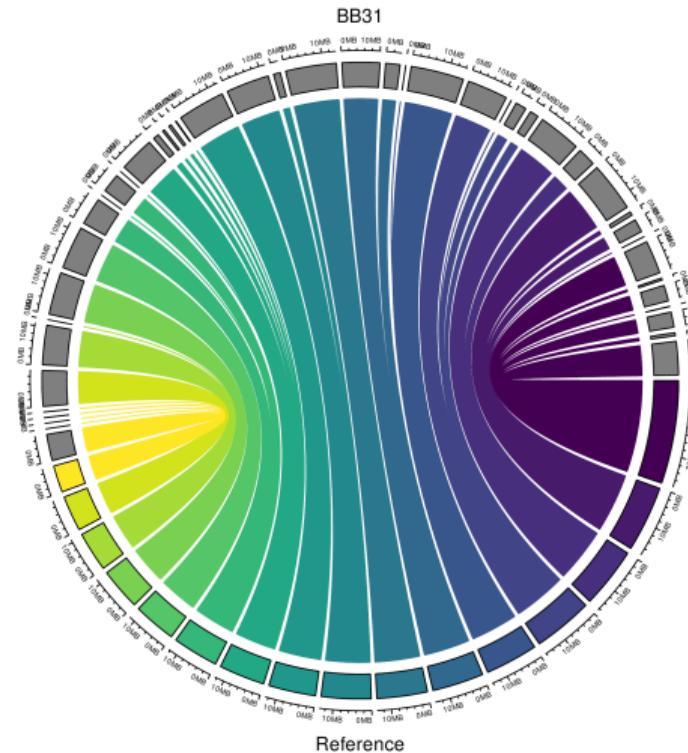
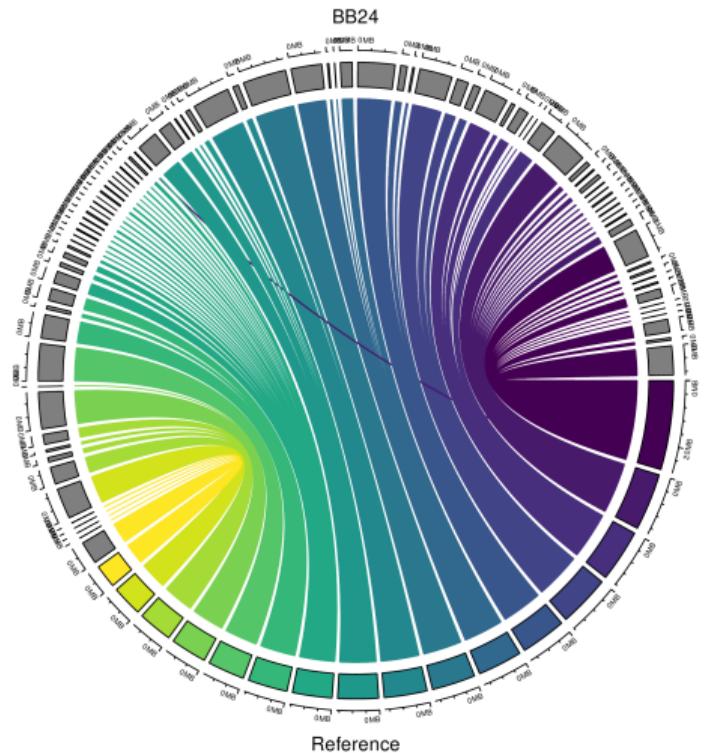
Metric	Amel_4.0	Amel_4.5	Amel_HAv3	Draft	Target
Contig N₅₀	40,692	45,822	5,382,476	> 10 Kb	> 1 Mb
Scaffold N₅₀	361,830	13,219,345	13,619,445	> 100 Kb	> 10 Mb
Gaps	8,202	11,181	51	< 10,000	< 1000
Completeness	🟡	🟡	🟡	> 80%	> 90%
Genes¹	95.00%	96.40%	97.70%	> 80%	> 90%
Mappability²	90.09%	90.22%	91.26%	> 70%	> 80%

1. BUSCO, *Benchmarking Universal Single-Copy Orthologs*
2. Illumina data from two individual honeybees from NZ

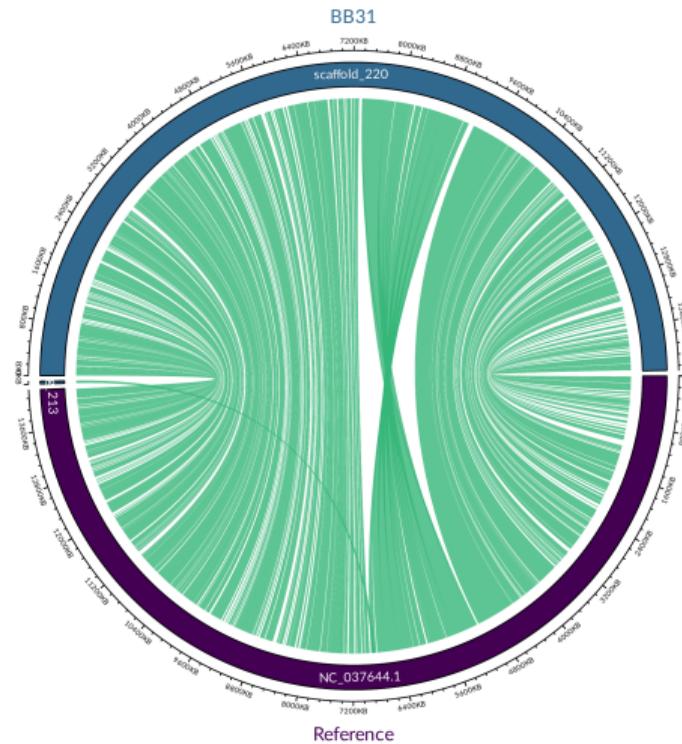
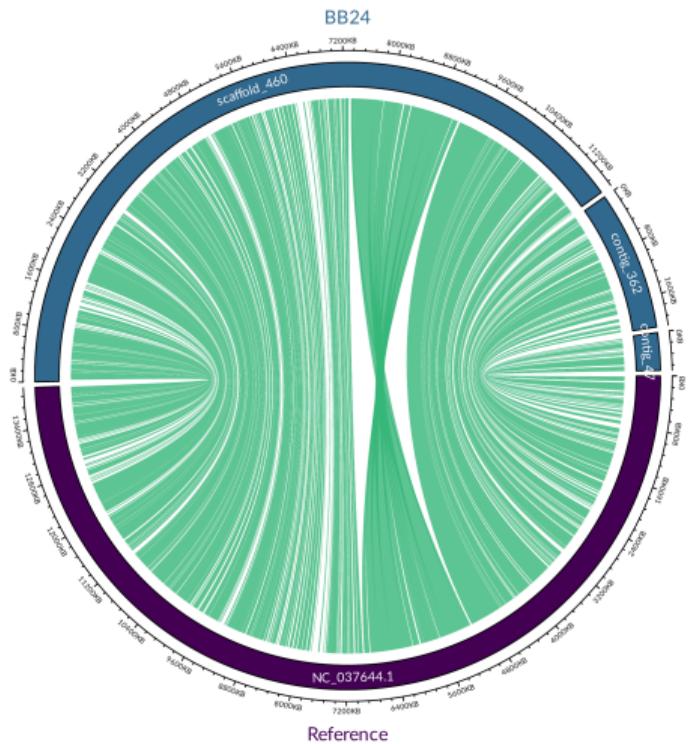
Is this a representative genome?



We can assemble local genomes to compare



Genomic rearrangements



Graphs and pangenomes

Method | [Open Access](#) | Published: 16 October 2020

The design and construction of reference pangenome graphs with minigraph

[Heng Li](#) , [Xiaowen Feng](#) & [Chong Chu](#)

[Genome Biology](#) 21, Article number: 265 (2020) | [Cite this article](#)

6269 Accesses | 9 Citations | 139 Altmetric | [Metrics](#)

Method | [Open Access](#) | Published: 12 February 2020

Genotyping structural variants in pangenome graphs using the vg toolkit

[Glenn Hickey](#), [David Heller](#), [Jean Monlong](#), [Jonas A. Sibbesen](#), [Jouni Sirén](#), [Jordan Eizenga](#), [Eric T. Dawson](#), [Erik Garrison](#), [Adam M. Novak](#) & [Benedict Paten](#) 

[Genome Biology](#) 21, Article number: 35 (2020) | [Cite this article](#)

6671 Accesses | 13 Citations | 16 Altmetric | [Metrics](#)