Variant detection and analysis tools

PRINCESS

- Mahmoud M, Doddapaneni H, Timp W, Sedlazeck FJ. PRINCESS: comprehensive detection of haplotype resolved SNVs, SVs, and methylation. *Genome Biology*. 2021;22(1). doi:https://doi.org/10.1186/s13059-021-02486-w
- https://github.com/MeHelmy/princess

Short variants with short reads

GATK (Genome Analysis Toolkit)

- McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*. 2010;20(9):1297-1303. doi:https://doi.org/10.1101/gr.107524.110
- https://gatk.broadinstitute.org/

DRAGEN (Dynamic Read Analysis for GENomics)

- https://www.illumina.com/products/by-type/informatics-products/dragen-secondary-analysis.html

DRAGEN-GATK

- https://gatk.broadinstitute.org/hc/en-us/articles/360045944831-DRAGEN-GATK

FreeBayes

- Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. *arXiv:1207.3907 [q-bio.GN].* 2012. doi: https://doi.org/10.48550/arXiv.1207.3907
- https://github.com/freebayes/freebayes

Samtools

- Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079. doi:https://doi.org/10.1093/bioinformatics/btp352
- Danecek P, Bonfield JK, Liddle J, et al. Twelve years of SAMtools and BCFtools. *GigaScience*. 2021;10(2). doi:https://doi.org/10.1093/gigascience/giab008

- https://github.com/samtools/samtools

Short variants with long reads

DeepVariant

- Poplin R, Chang PC, Alexander D, et al. A universal SNP and small-indel variant caller using deep neural networks. *Nature Biotechnology*. 2018;36(10):983-987. doi:https://doi.org/10.1038/nbt.4235
- https://github.com/google/deepvariant

Clair

- Luo R, Wong CL, Wong YS, et al. Exploring the limit of using a deep neural network on pileup data for germline variant calling. *Nature Machine Intelligence*. 2020;2(4):220-227. doi:https://doi.org/10.1038/s42256-020-0167-4
- https://github.com/HKU-BAL/Clair

Clair3

- Zheng Z, Li S, Su J, Leung A, Lam TW, Luo R. Symphonizing pileup and full-alignment for deep learning-based long-read variant calling. *Nature Computational Science*. 2022;2(12):797-803. doi:https://doi.org/10.1038/s43588-022-00387-x
- https://github.com/HKU-BAL/Clair3

Clair3-trio

- Su J, Zheng Z, Ahmed SS, Lam TW, Luo R. Clair3-trio: high-performance Nanopore long-read variant calling in family trios with trio-to-trio deep neural networks. *Briefings in Bioinformatics*. 2022;23(5). doi:https://doi.org/10.1093/bib/bbac301
- https://github.com/HKU-BAL/Clair3-Trio

Longshot

- Edge P, Bansal V. Longshot enables accurate variant calling in diploid genomes from single-molecule long read sequencing. *Nature Communications*. 2019;10(1). doi:https://doi.org/10.1038/s41467-019-12493-y
- https://github.com/pjedge/longshot

Medaka

- https://github.com/nanoporetech/medaka

Structural variants with short reads

Parliament2

- Zarate S, Carroll A, Mahmoud M, et al. Parliament2: Accurate structural variant calling at scale. *GigaScience*. 2020;9(12). doi:https://doi.org/10.1093/gigascience/giaa145
- https://github.com/fritzsedlazeck/parliament2

DELLY

- Rausch T, Zichner T, Schlattl A, Stütz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics*. 2012;28(18):i333-i339. doi:https://doi.org/10.1093/bioinformatics/bts378
- https://github.com/dellytools/delly

LUMPY

- Layer RM, Chiang C, Quinlan AR, Hall IM. LUMPY: a probabilistic framework for structural variant discovery. *Genome Biology*. 2014;15(6):R84. doi:https://doi.org/10.1186/gb-2014-15-6-r84
- https://github.com/arq5x/lumpy-sv

Manta

- Chen X, Schulz-Trieglaff O, Shaw R, et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics*. 2015;32(8):1220-1222. doi:https://doi.org/10.1093/bioinformatics/btv710
- https://github.com/Illumina/manta

Structural variants with long reads

Sniffles

- Sedlazeck FJ, Rescheneder P, Smolka M, et al. Accurate detection of complex structural variations using single-molecule sequencing. *Nature Methods*. 2018;15(6):461-468. doi:https://doi.org/10.1038/s41592-018-0001-7
- https://github.com/fritzsedlazeck/Sniffles

Sniffles2

- Smolka M, Paulin LF, Grochowski CM, et al. Comprehensive Structural Variant Detection: From Mosaic to Population-Level. *bioRxiv*. Published online April 5, 2022. doi:https://doi.org/10.1101/2022.04.04.487055
- https://github.com/fritzsedlazeck/Sniffles

cuteSV

- Jiang T, Liu Y, Jiang Y, et al. Long-read-based human genomic structural variation detection with cuteSV. *Genome Biology*. 2020;21(1). doi:https://doi.org/10.1186/s13059-020-02107-y
- Cao S, Jiang T, Liu Y, et al. Re-genotyping structural variants through an accurate force-calling method. *bioRxiv*. Published online September 9, 2022.; doi:https://doi.org/10.1101/2022.08.29.505534
- https://github.com/tjiangHIT/cuteSV

PBSV

- https://github.com/PacificBiosciences/pbsv

Dipcall

- Li H, Bloom JM, Farjoun Y, et al. A synthetic-diploid benchmark for accurate variant-calling evaluation. *Nature Methods*. 2018;15(8):595-597. doi:https://doi.org/10.1038/s41592-018-0054-7
- https://github.com/lh3/dipcall

Genome assembly and analysis tools

Short-read assemblies

SPAdes

- Bankevich A, Nurk S, Antipov D, et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*. 2012;19(5):455-477. doi:https://doi.org/10.1089/cmb.2012.0021
- Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes De Novo Assembler. *Current Protocols in Bioinformatics*. 2020;70(1). doi:https://doi.org/10.1002/cpbi.102
- https://github.com/ablab/spades

ABySS

- Simpson JT, Wong K, Jackman SD, et al. ABySS: a parallel assembler for short read sequence data. *Genome Research*. 2009;19(6):1117-1123. doi:https://doi.org/10.1101/gr.089532.108
- Jackman SD, Vandervalk BP, Mohamadi H, et al. ABySS 2.0: Resource-efficient assembly of large genomes using a Bloom filter. *Genome Research*. 2017;27(5): 768-777. doi:https://doi.org/doi:10.1101/gr.214346.116
- https://github.com/bcgsc/abyss

Velvet

- Zerbino DR, Birney E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*. 2008;18(5):821-829. doi:https://doi.org/10.1101/gr.074492.107
- Zerbino DR. Using the Velvet de novo Assembler for Short-Read Sequencing Technologies. *Current Protocols in Bioinformatics*. 2010;31(1). doi:https://doi.org/10.1002/0471250953.bi1105s31
- https://github.com/dzerbino/velvet

SOAPdenovo2

- Luo R, Liu B, Xie Y, et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience*. 2012;1(1). doi:https://doi.org/10.1186/2047-217x-1-18
- https://github.com/aquaskyline/SOAPdenovo2

Long-read assemblies

Canu

- Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptivek-mer weighting and repeat separation. *Genome Research*. 2017;27(5):722-736. doi:https://doi.org/10.1101/gr.215087.116
- Koren S, Rhie A, Walenz BP, et al. De novo assembly of haplotype-resolved genomes with trio binning. *Nature Biotechnology.* 2018;36(12):1174-1182. doi:https://doi.org/10.1038/nbt.4277
- https://github.com/marbl/canu

Shasta

- Shafin K, Pesout T, Lorig-Roach R, et al. Nanopore sequencing and the Shasta toolkit enable efficient de novo assembly of eleven human genomes. *Nature Biotechnology*. 2020;38(9). doi:https://doi.org/10.1038/s41587-020-0503-6
- https://github.com/paoloshasta/shasta

Flye

- Lin Y, Yuan J, Kolmogorov M, Shen MW, Chaisson M, Pevzner PA. Assembly of long error-prone reads using de Bruijn graphs. *Proceedings of the National Academy of Sciences*. 2016;113(52):E8396-E8405. doi:https://doi.org/10.1073/pnas.1604560113
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology*. 2019;37(5):540-546. doi:https://doi.org/10.1038/s41587-019-0072-8
- Kolmogorov M, Bickhart DM, Behsaz B, et al. metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nature Methods*. 2020;17(11):1103-1110. doi:https://doi.org/10.1038/s41592-020-00971-x
- https://github.com/fenderglass/Flye

wtdbg2

- Ruan J, Li H. Fast and accurate long-read assembly with wtdbg2. *Nature Methods*. 2019;17(2). doi:https://doi.org/10.1038/s41592-019-0669-3
- https://github.com/ruanjue/wtdbg2