Supplementary material

Effects of Land Use and Landscape Fragmentation on the Genetic Composition of Arthropods

These are the references for *Genetic Load in Populations of Collembola (Springtails) in a Fragmented Landscape*

References

Bao, Weidong, Kenji K. Kojima, and Oleksiy Kohany (June 2015). "Repbase Update, a database of repetitive elements in eukaryotic genomes". In: *Mobile DNA* 6 (11). ISSN: 17598753. DOI: 10.1186/s13100-015-0041-9.

Brůna, Tomáš, Alexandre Lomsadze, and Mark Borodovsky (2024). "A new gene finding tool GeneMark-ETP significantly improves the accuracy of automatic annotation of large eukaryotic genomes". In: *bioRxiv*. DOI: 10.1101/2023.01.13.524024.

Brůna, Tomáš et al. (Mar. 2021). "BRAKER2: Automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database". In: *NAR Genomics and Bioinformatics* 3 (1), pp. 1–11. ISSN: 26319268. DOI: 10.1093/nargab/lqaa108.

Buchfink, Benjamin, Chao Xie, and Daniel H. Huson (Jan. 2015). *Fast and sensitive protein alignment using DIAMOND*. DOI: 10.1038/nmeth.3176.

Cheng, Haoyu et al. (Feb. 2021). "Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm". In: *Nature Methods* 18 (2), pp. 170–175. ISSN: 15487105. DOI: 10.1038/s41592-020-01056-5.

Cheng, Haoyu et al. (Sept. 2022). "Haplotype-resolved assembly of diploid genomes without parental data". In: *Nature Biotechnology* 40 (9), pp. 1332–1335. ISSN: 15461696. DOI: 10.1038/s41587-022-01261-x.

Cingolani, Pablo et al. (2012). "A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3". In: *Fly* 6 (2), pp. 80–92. ISSN: 19336942. DOI: 10.4161/fly.19695.

Dainat, Jacques (2023). *AGAT: Another Gff Analysis Toolkit to handle annotations in any GTF/GFF format (Version v1.2.0)*. DOI: 10.5281/zenodo.3552717.

Danecek, Petr et al. (Feb. 2021). "Twelve years of SAMtools and BCFtools". In: *GigaScience* 10 (2). ISSN: 2047217X. DOI: 10.1093/gigascience/giab008.

Dudchenko, Olga et al. (Apr. 2017). "De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds". In: *Science* 356 (6333), pp. 92–95. ISSN: 10959203. DOI: 10.1126/science.aal3327.

Durand, Neva C. et al. (July 2016a). "Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom". In: *Cell Systems* 3 (1), pp. 99–101. ISSN: 24054720. DOI: 10.1016/j.cels.2015.07.012. URL: http://dx.doi.org/10.1016/j.cels.2015.07.012.

Durand, Neva C. et al. (July 2016b). "Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments". In: *Cell Systems* 3 (1), pp. 95–98. ISSN: 24054720. DOI: 10.1016/j.cels.2016.07.002.

Flynn, Jullien M. et al. (Apr. 2020). "RepeatModeler2 for automated genomic discovery of transposable element families". In: *PNAS* 117 (17), pp. 9451–9457. URL: https://doi.org/10.1073/pnas.1921046117.

Gabriel, Lars et al. (Dec. 2021). "TSEBRA: transcript selector for BRAKER". In: *BMC Bioinformatics* 22. ISSN: 14712105. DOI: 10.1186/s12859-021-04482-0.

Gabriel, Lars et al. (May 2024). "BRAKER3: Fully automated genome annotation using RNA-seq and protein evidence with GeneMark-ETP, AUGUSTUS, and TSEBRA". In: *Genome Research* 34, pp. 769–777. ISSN: 1088-9051. DOI: 10.1101/gr.278090.123.

García-Alcalde, Fernando et al. (Oct. 2012). "Qualimap: Evaluating next-generation sequencing alignment data". In: *Bioinformatics* 28 (20), pp. 2678–2679. ISSN: 13674803. DOI: 10.1093/bioinformatics/bts503.

Garrison, Erik and Gabor Marth (July 2012). "Haplotype-based variant detection from short-read sequencing". In: *arXiv*. DOI: 10.48550/arXiv.1207.3907. URL: http://arxiv.org/abs/1207.3907https://doi.org/10.48550/arXiv.1207.3907.

Gotoh, Osamu (May 2008). "A space-efficient and accurate method for mapping and aligning cDNA sequences onto genomic sequence". In: *Nucleic Acids Research* 36 (8), pp. 2630–2638. ISSN: 03051048. DOI: 10.1093/nar/gkn105.

Guan, Dengfeng et al. (May 2020). "Identifying and removing haplotypic duplication in primary genome assemblies". In: *Bioinformatics* 36 (9), pp. 2896–2898. ISSN: 14602059. DOI: 10.1093/bioinformatics/btaa025.

Hoff, Katharina J. et al. (Mar. 2016). "BRAKER1: Unsupervised RNA-Seq-based genome annotation with GeneMark-ET and AUGUSTUS". In: *Bioinformatics* 32 (5), pp. 767–769. ISSN: 14602059. DOI: 10.1093/bioinformatics/btv661.

[Internet], Assembly (Aug. 2016). *GenBank assembly accession: $GCA_0$01718145.1, $BioSampleID$: $SAMN$04014362, $Orchesellacincta$, haploid*. URL: https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_001718145.1.

Iwata, Hiroaki and Osamu Gotoh (Nov. 2012). "Benchmarking spliced alignment programs including Spaln2, an extended version of Spaln that incorporates additional species-specific features". In: *Nucleic Acids Research* 40 (20), e161. ISSN: 03051048. DOI: 10.1093/nar/gks708.

Li, Heng (Mar. 2013). "Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM". In: URL: http://arxiv.org/abs/1303.3997.

— (Sept. 2018). "Minimap2: Pairwise alignment for nucleotide sequences". In: *Bioinformatics* 34 (18), pp. 3094–3100. ISSN: 14602059. DOI: 10.1093/bioinformatics/bty191.

— (Dec. 2021). "New strategies to improve minimap2 alignment accuracy". In: *Bioinformatics* 37 (23), pp. 4572–4574. ISSN: 14602059. DOI: 10.1093/bioinformatics/btab705.

Manni, Mosè et al. (Oct. 2021). "BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes". In: *Molecular Biology and Evolution* 38 (10), pp. 4647–4654. ISSN: 15371719. DOI: 10.1093/molbev/msab199.

Mathur, Samarth et al. (2023). "Functional genomic diversity is correlated with neutral genomic diversity in populations of an endangered rattlesnake". In: *Proceedings of the National Academy of Sciences of the United States of America* 120 (43). ISSN: 10916490. DOI: 10.1073/pnas.2303043120.

Pertea, Mihaela and Geo Pertea (2020). "GFF Utilities: GffRead and GffCompare". In: *F1000Research* 9. ISSN: 1759796X. DOI: 10.12688/f1000research.23297.1.

Quinlan, Aaron R. and Ira M. Hall (Jan. 2010). "BEDTools: A flexible suite of utilities for comparing genomic features". In: *Bioinformatics* 26 (6), pp. 841–842. ISSN: 13674803. DOI: 10.1093/bioinformatics/btq033.

Schubert, Mikkel, Stinus Lindgreen, and Ludovic Orlando (Feb. 2016). "AdapterRemoval v2: Rapid adapter trimming, identification, and read merging". In: *BMC Research Notes* 9. ISSN: 17560500. DOI: 10.1186/s13104-016-1900-2.

Sim, Sheina B. et al. (Dec. 2022). "HiFiAdapterFilt, a memory efficient read processing pipeline, prevents occurrence of adapter sequence in PacBio HiFi reads and their negative impacts on genome assembly". In: *BMC Genomics* 23 (1). ISSN: 14712164. DOI: 10.1186/s12864-022-08375-1.

Smit, A. F. A., R. Hubley, and P. Green (2013). *RepeatMasker Open-4.0. 2013-2015*. URL: http://repeatmasker.org.

Stanke, Mario et al. (Feb. 2006). "Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources". In: *BMC Bioinformatics* 7. ISSN: 14712105. DOI: 10.1186/1471-2105-7-62.

Stanke, Mario et al. (Jan. 2008). "Using native and syntenically mapped cDNA alignments to improve de novo gene finding". In: *Bioinformatics* 24 (5), pp. 637–644. ISSN: 13674803. DOI: 10.1093/bioinformatics/btn013.

Stoletzki, Nina and Adam Eyre-Walker (Jan. 2011). "Estimation of the neutrality index". In: *Molecular Biology and Evolution* 28 (1), pp. 63–70. ISSN: 07374038. DOI: 10.1093/molbev/msq249.