GCBN

Service and Consulting for Annotation of Complex Crop (Pan) Genomes

Heidrun Gundlach[1], Thomas Lux[1], Georg Haberer[1], Manuel Spannagl[1], Klaus FX Mayer[1]
Helmholtz Zentrum München Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH)

Contract Nr. W-de.NBI-xxx

Short description of the project

The annotation of large and complex (crop) plant genomes remains a challenging task, particularly with the rapid integration of newly emerging technologies in this dynamic field. Leveraging extensive expertise at the forefront of reference and pan-genomics, our HMGU team offers comprehensive genome annotation services, including the provision of high-quality basic and processed data, as well as expert consulting tailored to the needs of the plant research community. We provide customized in-house training programs for early-career scientists, empowering them to develop and implement their own annotation workflows. We collaborate closely with our GCBN partners, IPK and FZJ, on data generation and the delivery of FAIR-compliant data sets. Additionally we support the activities of ELIXIR as a member of the Plant Science Community contributing to ELIXIR's "Biodiversity, Food Security, & Pathogens" priority area.

de.NBI services

The next scale of plant genome annotation: pan genomes with 30 to >100 varieties per species

pasta wheat

barley and wild barley

hop

50 x 11 Gbp = 550 Gbp
to annotate

fescue grass

faba bean

bread wheat and wild ancestors

Gene annotation strategy :

Step 1

de-novo gene predictions

Line-specific gene structures and gene models, using extensive transcriptome data

Step 2

consolidated gene projections

fast&relieable assessment of gene content

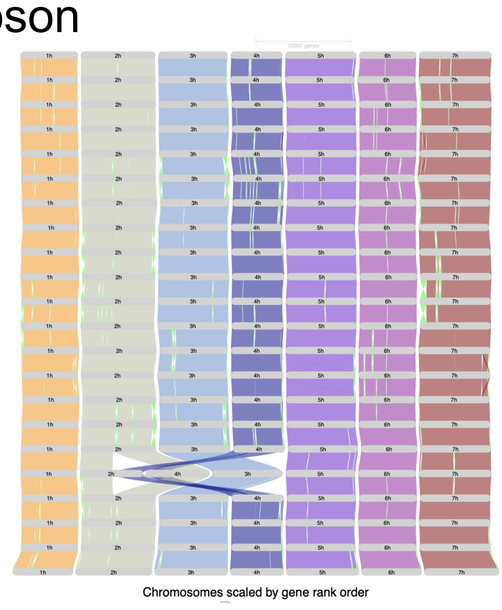
scalable for the next 100 genomes

General information on the project

transfer agreement not yet signed by HMGU
intended funding 0.75 FTE
matching fund: 0.75 FTE

Progress report

- enhanced gene and transposon annotation pipelines
- developed a gene projection workflow for pan-genomes
- provided annotations for 20 single genomes and 4 pan-genome projects



- active member of ELIXIR Plant Science community
- participation in de.NBI & ELIXIR Hackathons
- ELIXIR projects: "Increasing plant data findability and reuse beyond ELIXIR"
  "E-PAN: Enhancing pan-genome analysis in plants"

de.NBI **Training** and education

- tailored supervision for 7 trainees, each staying 4-12 months
- trainees worked on their own datasets
- they left with tangible results and the skills to build own bioinformatics pipelines

Publications

Jayakodi, M., Lu, Q., Pidon, H. et al. **Structural variation in the pangenome of wild and domesticated barley.** Nature (2024).

Franzisky, BL., Mueller, HM., Du, B, Lux, T. et al. **Date palm diverts organic solutes for root osmotic adjustment and protects leaves from oxidative damage in early drought acclimation.** Journal of Experimental Botany, erae456 (2024)

Avni, R., Kamal, N., Bitz, L., Jellen, EN. et al. **A pangenome and pantranscriptome of hexaploid oat**. preprint, BioXiv (2024).

White, B., Lux, T., Rusholme-Pilcher, R., Juhász, A. **De novo annotation of the wheat pan-genome reveals complexity and diversity of the hexaploid wheat pan-transcriptome.** Preprint, BioXiv (2024).

Garg, V., Bohra, A., Mascher, M. et al. **Unlocking plant genetics with telomere-to-telomere genome assemblies.** Nat Genet 56, 1788–1799 (2024).

Coombes, B., Lux, T., Akhunov, E. et al. **Introgressions lead to reference bias in wheat RNA-seq analysis.** BMC Biol 22, 56 (2024).

Mascher, M., Feng JW, Pidon H. et al. **A haplotype-resolved pangenome of the barley wild relative Hordeum bulbosum.** Preprint, Research Square (2024).

Castellani, M., Zhang, M., Thangavel, G. et al. **Meiotic recombination dynamics in plants with repeat-based holocentromeres shed light on the primary drivers of crossover patterning.** Nat. Plants 10, 423–438 (2024).

Jellen EN., Wight CP., Spannagl M., Blake VC. et al. **A uniform gene and chromosome nomenclature system for oat. (Avena spp.).** Crop & Pasture Science 75, CP23247. (2024).

Robbie W., Wenbin G. Miriam S.et al. **A barley pan-transcriptome reveals layers of genotype-dependent transcriptional complexity.** Preprint, Research Square (2024).

Mueller, HM., Franzisky, BL., Messerer, M., Du, B. et al. **Integrative multi-omics analyses of date palm (Phoenix dactylifera) roots and leaves reveal how the halophyte land plant copes with sea water.** The Plant Genome. 17, e20372 (2023).

Devos, K.M., Qi, P., Bahri, B.A. et al. **Genome analyses reveal population structure and a purple stigma color gene candidate in finger millet.** Nat Commun 14, 3694 (2023).

Njaci, I., Waweru, B., Kamal, N. et al. **Chromosome-level genome assembly and population genomic resource to accelerate orphan crop lablab breeding.** Nat Commun 14, 1915 (2023).

Jayakodi, M., Golicz, A.A., Kreplak, J. et al. **The giant diploid faba genome unlocks variation in a global protein crop.** Nature 615, 652–659 (2023).

Bamba, M.; Shirasawa, K.; Isobe, S.; Kamal, N. et al. **Plant Genomics.** Plant Omics: Advances in Big Data Biology, Cabi Publishing pp. 1 (2023).

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung