GCBN

FZJ-IBG-4’s Contributions to the de.NBI and ELIXIR-DE Networks

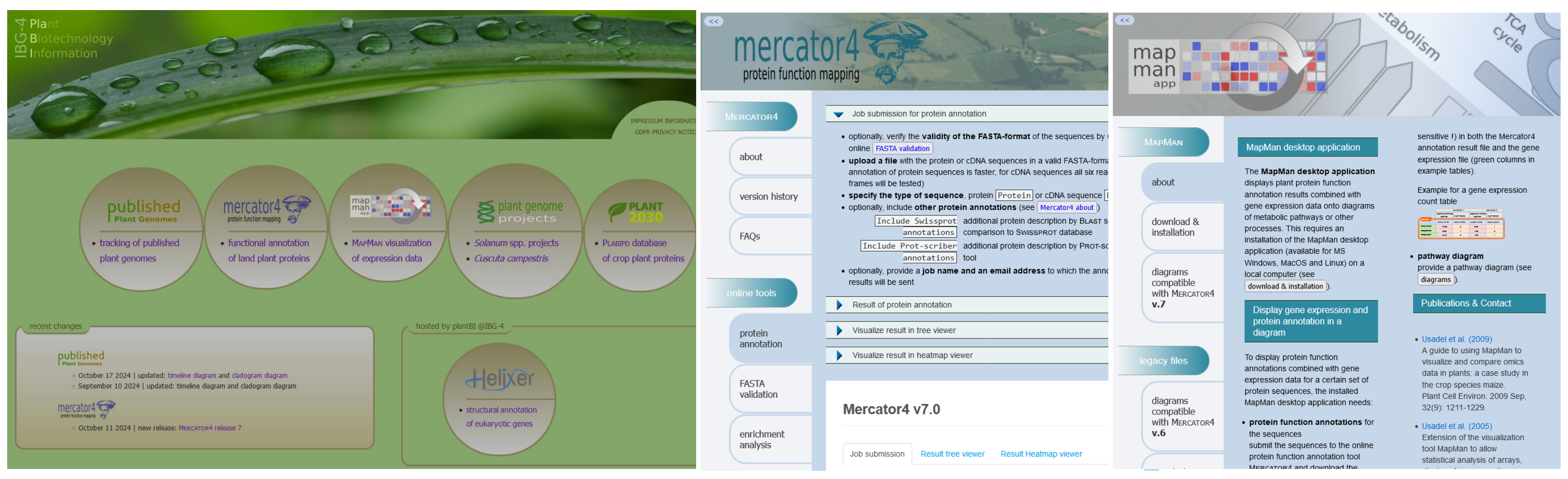
Sebastian Beier[1], Marie Bolger[1], Rainer Schwacke[1], Felicitas Kindel[1], Björn Usadel[1]

1 Institute of Bio- and Geosciences (IBG-4: Bioinformatics), Bioeconomy Science Center (BioSC), CEPLAS, Forschungszentrum Jülich GmbH, Jülich, Germany

Short description of the project

- **Contribution to the GCBN Service Unit within de.NBI**: Providing bioinformatics Services (see below), offer consulting and training in gene finding, annotation, pangenome analysis, genome studies and comparative genomics.
- **Integration with International Research**: Supporting integration with high-throughput plant phenotyping platforms and aligning with EU-wide projects and research infrastructures to enhance research capabilities.
- **Event Organization and Community Networking**: Coordinating and participating in de.NBI and ELIXIR events and fostering connections across the Plant Community to strengthen collaboration and knowledge sharing.

de.NBI services



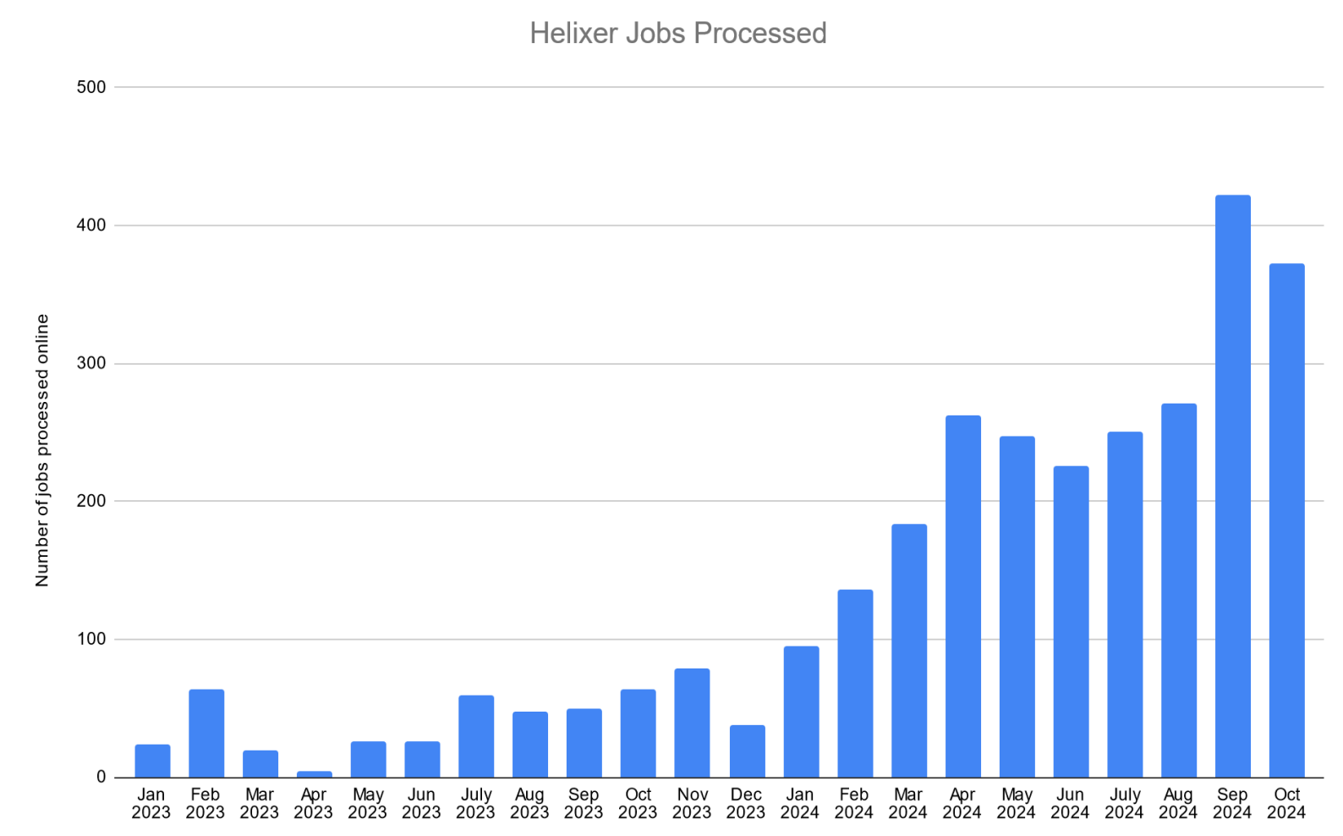
MapMan: A visualization tool that enables researchers to map and interpret gene expression data within biological pathways and metabolic processes in plants.

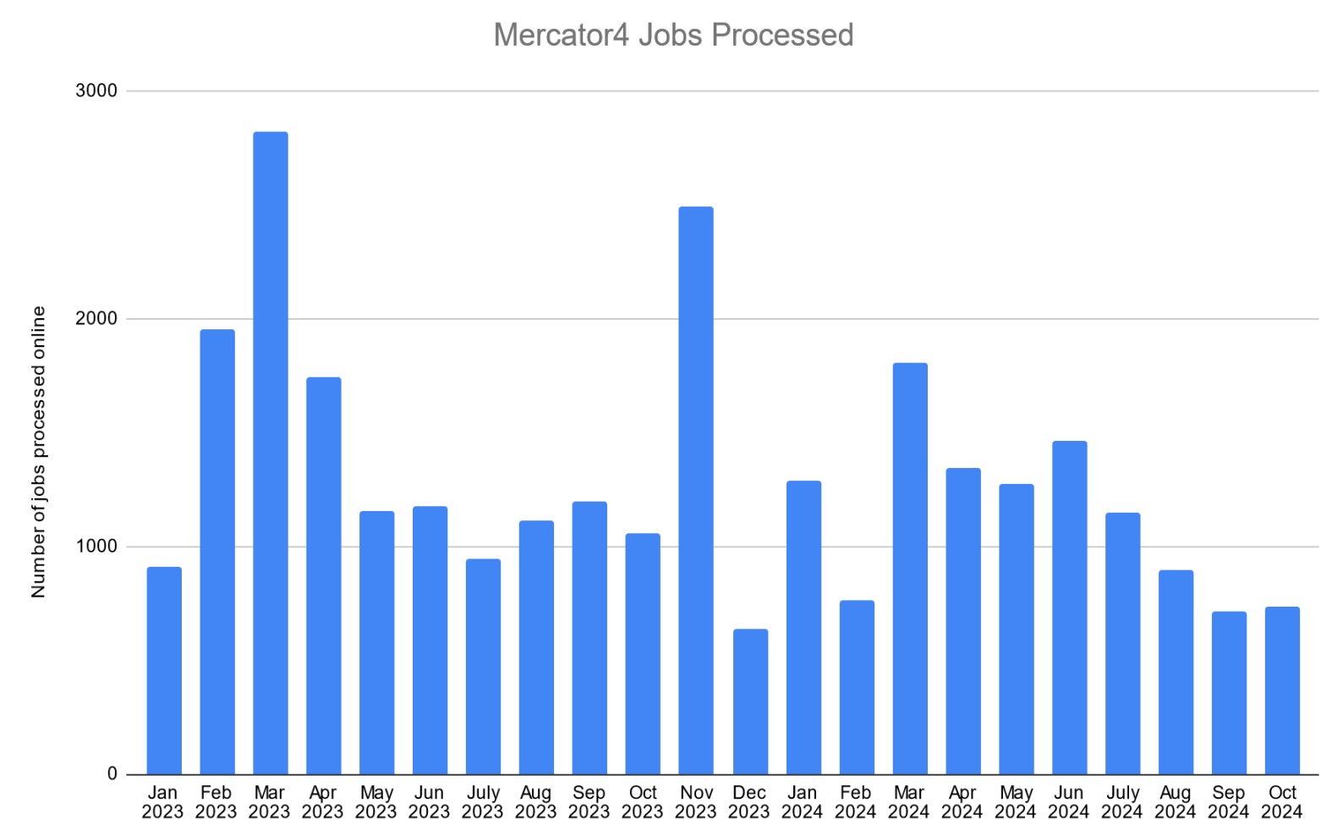
Trimmomatic: A flexible tool for trimming and cleaning sequencing reads, optimizing data quality for downstream analyses in bioinformatics workflows.

Mercator4: A tool for functional annotation of plant proteins, allowing for detailed categorization based on specific biological processes and comparative genomics.

PlabiPD: A plant bioinformatics platform providing access to curated data and tools for the analysis of plant genomes and transcriptomes.

Helixer: A deep learning-based tool that accurately predicts structural genes in nucleotide sequences without the need for RNA-Seq data.





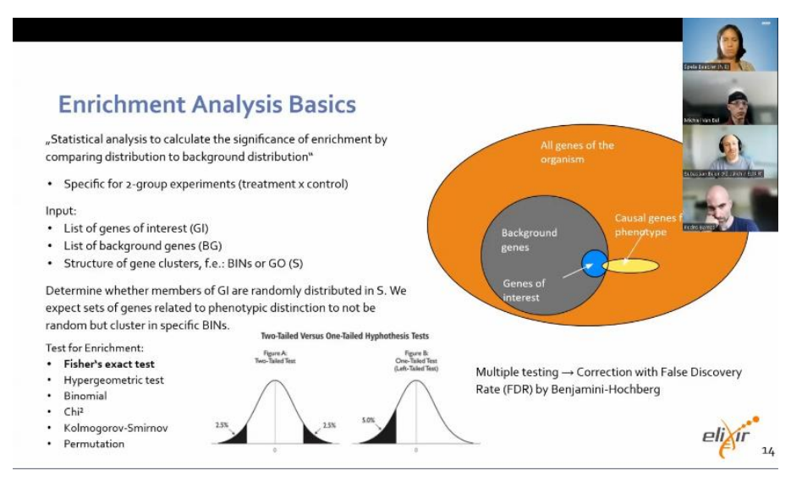
Progress report

- **Service Development**: IBG-4 contributed to the enhancement of de.NBI services, with notable upgrades to **Mercator4** (versions 6 and 7), introducing features such as the protein length distribution plot, a new Plant Organogenesis category, and hundreds of new protein categories and context nodes. An **Enrichment App** was also added to Mercator4, enabling Over-Representation Analysis (ORA) and Under-Representation Analysis (URA) for gene sets annotated by Mercator4. Additionally, the release of **Helixer** introduced a deep learning tool for structural genome annotation without the need for RNA-Seq data.
- **KPI Monitoring**: Consistent growth in citations and downloads across de.NBI tools reflects strong user engagement. We are transitioning from Google Analytics (GA) to **Matomo** to improve privacy; a brief KPI data gap occurred during an unapproved switch between GA versions.
- **ELIXIR Engagement**: IBG-4 has been actively involved in ELIXIR, with Sebastian Beier serving as the GCBN Officer and Co-lead of the Plant Sciences Community. IBG-4 also leads the new Commissioned Service (CoS) **HARVEST**, aimed at enhancing plant research FAIRification through ELIXIR data standards, and is a partner in **E-PAN**, a project focused on advancing plant pan-genome analysis within ELIXIR.

de.NBI Training and education

Since 2023, our de.NBI and ELIXIR trainings have provided essential skills in bioinformatics and RDM. Notable events include:

- **de.NBI Spring School 2023**: Focused on FAIR data management at IPK Gatersleben with hands-on sessions.
- **de.NBI Summer School 2023**: Offered an overview of high-throughput data processing and analysis for life sciences, held in Bielefeld.
- **ELIXIR INCREASING Virtual Workshop 2023**: Showcased plant tool bundles for plant data analysis and end point repository submission.
- **Field-Ready Plant Bioinformatics**: Training on genome annotation tools, hosted in Jülich in October 2024.



General information on the project

- **2 FTE** paid from de.NBI grant
- 2 more FTE involved but not paid by de.NBI grant

Publications

Improving Metadata Collection and Aggregation in Plant Phenotyping Experiments with MIAPPE Wizard and DataPLANT. Arend D, **Beier S**, Brilhaus D, **Dörpholz H**, Feser M, Frey K, König P, Maus O, Psaroudakis D, Rodrigues CM, Schrader A, **Senger E**, Weil HL. BioHackrXiv; 2023. DOI: 10.37044/osf.io/ekhdw.

DivBrowse-interactive visualization and exploratory data analysis of variant call matrices. König P, **Beier S**, Mascher M, Stein N, Lange M, Scholz U Gigascience. 2022 Dec;12:giad025. DOI: 10.1093/gigascience/giad025. PMID: 37083938; PMCID: PMC10120423.

Plant science data integration, from building community standards to defining a consistent data lifecycle. In: H.F. Williamson & S. Leonelli (Eds.) Towards responsible plant data linkage: data challenges for agricultural research and development. Pommier C, Coppens F, Ćwiek-Kupczyńska H, Faria D, **Beier S**, Miguel C, Michotey C, D’anna F, Owen S, Gruden K. Cham: Springer International Publishing, pp. 149–160. Available from: https://doi.org/10.1007/978-3-031-13276-6_8

BioHackEU23 report: Enabling continuous RDM using Annotated Research Contexts with RO-Crate profiles for ISA. **Beier S**, Mühlhaus T, Pommier C, Owen S, Brilhaus D, Weil HL, Wetzels F, Chait G, Arend D, Feser M, Doniparthi G, Bauer J, Gundersen S, Vazquez P. BioHackrXiv; 2024. DOI: 10.37044/osf.io/7y2jh.

A Solanum lycopersicoides reference genome facilitates insights into tomato specialized metabolism and immunity. Powell AF, Feder A, Li J, **Schmidt MH**, Courtney L, Alseekh S, Jobson EM, Vogel A, Xu Y, Lyon D, **Dumschott K**, McHale M, Sulpice R, Bao K, Lal R, Duhan A, **Hallab A**, Denton AK, **Bolger ME**, Fernie AR, Hind SR, Mueller LA, Martin GB, Fei Z, Martin C, Giovannoni JJ, Strickler SR, **Usadel B**. Plant J. 2022 Jun;110(6):1791-1810. doi: 10.1111/tpj.15770. Epub 2022 May 18. PMID: 35411592

Phased, chromosome-scale genome assemblies of tetraploid potato reveal a complex genome, transcriptome, and predicted proteome landscape underpinning genetic diversity. Hoopes G, Meng X, Hamilton JP, Achakkagari SR, de Alves Freitas Guesdes F, **Bolger ME**, Coombs JJ, Esselink D, Kaiser NR, Kodde L, Kyriakidou M, Lavrijssen B, van Lieshout N, Shereda R, Tuttle HK, Vaillancourt B, Wood JC, de Boer JM, Bornowski N, Bourke P, Douches D, van Eck HJ, Ellis D, Feldman MJ, Gardner KM, Hopman JCP, Jiang J, De Jong WS, Kuhl JC, Novy RG, Oome S, Sathuvalli V, Tan EH, Ursum RA, Vales MI, Vining K, Visser RGF, Vossen J, Yencho GC, Anglin NL, Bachem CWB, Endelman JB, Shannon LM, Strömvik MV, Tai HH, **Usadel B**, Buell CR, Finkers R. Mol Plant. 2022 Mar 7;15(3):520-536. doi: 10.1016/j.molp.2022.01.003. Epub 2022 Jan 10. PMID: 35026436

MapMan Visualization of RNA-Seq Data Using Mercator4 Functional Annotations. **Bolger M**, **Schwacke R**, **Usadel B**. Methods Mol Biol. 2021;2354:195-212. doi: 10.1007/978-1-0716-1609-3_9. PMID: 34448161

Taken to extremes: Loss of plastid rpl32 in Streptophyta and Cuscuta's unconventional solution for its replacement. Fischer K, Valentin Jordbræk S, Olsen S, Bockwoldt M, **Schwacke R**, **Usadel B**, Krause K. Mol Phylogenet Evol. 2024 Nov 22:108243. doi: 10.1016/j.ympev.2024.108243. Epub ahead of print. PMID: 39581358.

A full genome assembly reveals drought stress effects on gene expression and metabolite profiles in blackcurrant (Ribes nigrum L.). **Ziegler FMR**, Rosenthal V, Vallarino JG, **Genzel F**, **Spettmann S**, Seliga L, Keller-Przybyłkowicz S, Munnes L, Sønsteby A, Osorio S, **Usadel B**. Horticulture Research. 2024. uhae313, doi: 10.1093/hr/uhae313

In-depth exploration of the genomic diversity in tea varieties based on a newly constructed pangenome of Camellia sinensis. Tariq A, Meng M, Jiang X, **Bolger A**, **Beier S**, Buchmann J, Fernie AR, Wen W, **Usadel B**. The Plant Journal : for Cell and Molecular Biology. 2024 Aug;119(4):2096-2115. DOI: 10.1111/tpj.16874. PMID: 38872506.

www.denbi.de

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung