

2026-27 HIVE Summer Internship Project

Interactive visualisation of complex genomic variation across crop genomes

Primary Academic Supervisor

Dr. Ulduz Vafadarshamasbi

Project Background

Modern crop genomics increasingly uses multiple genome assemblies rather than a single reference genome. These resources reveal extensive genomic diversity within crop species, including differences in genome structure, gene content and other forms of genetic variation. However, comparing complex genomic regions across multiple individuals remains challenging because information is distributed across genomes, coordinates and biological data layers.

This project will investigate how advanced interactive visualisation can improve the exploration and interpretation of complex comparative genomic data. Using publicly available and published datasets from a target crop, the student will examine existing approaches, compare alternative visual representations, and develop an interactive prototype for exploring genomic regions across multiple genome assemblies.

The project will provide the student with hands-on experience in comparative genomics, data analysis and advanced visualisation while investigating how HIVE facilities can support the exploration and communication of information-dense genomic datasets.

Project Description, Expected Outputs, Possible Stretch Goals

Aim:

Investigate and develop interactive visualisation approaches for exploring complex genomic variation across multiple genome assemblies of a target crop.

The student will use published genomic datasets to evaluate ways of representing genomic regions, genes and selected categories of variation, and will develop and test a prototype using HIVE facilities.

Expected outputs:

- Review and comparison of relevant genomic visualisation approaches.
- Curated, visualisation-ready public genomic data.
- Interactive prototype for comparing genomic regions across multiple assemblies.
- Functions such as zooming, filtering, feature selection and annotation display.
- Evaluation and refinement of the prototype on a HIVE display.
- Final research report and documentation.

Possible stretch goals include:

- Customised linked views using D3.js,
- Alternative visual encodings, or
- Exploration of a 3D representation where it provides a clear research benefit.

Potential tools include R, Python, Shiny, Plotly, BEDTools and minimap2, with the final selection depending on the visualisation approach developed during the project.

Links to background reading and any relevant recent work in the field

Suggested references from background reading:

- Bayer et al. (2020), Plant pan-genomes are the new reference, Nature Plants.
- Durant et al. (2021), Panache: a web browser-based viewer for linearized pangenomes, Bioinformatics.
- Technological Development and Advances for Constructing and Analyzing Plant Pangenomes (2024), Genome Biology and Evolution.
- PangyPlot: multi-scale interactive visualization of pangenome variation graphs (2026), Bioinformatics.

- Pretzel: interactive visualisation and integration of genetic and genomic datasets, Pretzel Documentation

What type of visualisation will the student develop or produce?

The student will investigate and develop an interactive comparative-genomics visualisation for exploring selected genomic regions across multiple genome assemblies. Alternative visual representations will be evaluated, with the final prototype supporting functions such as zooming, filtering, feature selection and annotation display. The visualisation will be tested and optimised for a HIVE display.

How will the visualisation contribute to your research outcomes?

The project will evaluate how interactive, large-format visualisation can improve exploration of complex multi-genome data. By integrating genomic regions, genes and selected variation into coordinated views, the prototype will help researchers recognise conserved and variable patterns, compare alternative representations, and establish a reusable visual approach for comparative genomics research.

If the project is successful, where would you consider publishing the results?

Potential outputs include a bioinformatics methods or software paper, or a conference presentation or poster, and potential incorporation of the visualisation into a broader crop-genomics publication. Suitable publication venues or relevant plant genomics and bioinformatics conferences will depend on the outcomes of the project.

Draft Project Timeline:

Week 1

Nov 16 - Full day HIVE induction

Nov 17 - Area and Project Induction with Primary supervisor

Develop project plan with HIVE and academic team

Literature and relevant project review

Project Familiarisation

Week 2

Explore and quality-check the published dataset. Review existing genome browsers and comparative visualisation approaches. Develop initial sketches/wireframes and identify candidate visual representations.

Week 3

Prepare visualisation-ready genomic data using R/Python and relevant bioinformatics tools. Develop and compare initial interactive visualisation approaches.

Week 4

Refine the most promising comparative visualisation approach.

Incorporate genomic regions, genes and selected variation layers. Review the prototype with supervisors.

Week 5

Implement interactive functions such as zooming, filtering, feature selection and annotation display. Evaluate usability and visual clarity using representative published examples. Complete a functional prototype.

Week 6

Extend and test the prototype across additional published examples.

Assess performance, readability and alternative visual representations for information-dense genomic data.

Week 7

Test the visualisation using the selected HIVE display. Evaluate its effectiveness for exploring complex genomic information, gather feedback and identify improvements.

Week 8

Refine the prototype based on testing and feedback and complete the core HIVE-ready visualisation. If time permits, explore an advanced feature such as customised D3.js linked views or an alternative visual representation.

Week 9

Focus on report writing and presentation (including demo) preparation

Week 10

Focus on report writing and presentation (including demo) preparation

5th Feb Final Showcase Day and final report due

Student Experience and Supervision:

How often will you meet with the student over the 10-week period?

At least one formal meeting per week with the primary academic supervisor, with additional short technical/check-in meetings as required. HIVE staff engagement will be scheduled at key design and display-testing milestones.

Your work desk location and the location of student desk

Building 303. A student desk will be provided nearby, with the exact location to be confirmed.

Student Attributes:

Please indicate any preference for student's academic discipline or field of study

Bioinformatics, Computational Biology, Data Science, Computer Science, or a related discipline. Prior crop-genomics experience is not required.

What competencies are required to start this project

Beginner - AI Integrations (build tools to assist with custom workflows)

Intermediate - Programming in R, Python or JavaScript

Intermediate - Interactive data visualisation / web-based visualisation

Do you have any other student attributes you think are important to the project?

Curiosity about scientific data, willingness to learn basic genomics concepts, attention to data interpretation, clear communication, and interest in interactive data visualisation. The student should be comfortable learning new tools and iterating on a prototype based on researcher feedback.