Philipp **Bayer**

RESEARCH OFFICER, AI

OceanOmics, Minderoo Foundation

✉ pbayer@minderoo.org | ⌨ philippbayer | 🔗 philippbayer | 🐦 philippbayer

About me

I work at the intersection of genomics, data science, machine learning and AI, high-performance computing, teaching, and environmental science, drawing on more than a decade of experience at the cutting edge of genomics to push what is possible in environmental and DNA data analysis. I have co-supervised more than twenty PhD and MSc students to completion and am now focusing on building a team of AI experts in eDNA.

Education

2012 to 2016	**PhD**, University of Queensland, Brisbane, Australia Applied bioinformatics PhD in the Edwards group. Developed computational pipeline SkimGBS for higher resolution genotyping. Worked extensively with industry (Bayer CropScience/BASF).
2010 to 2012	**Master of IT**, Bond University, Gold Coast, Australia Focused on coding and business IT. 5x Top of class, 3x Vice-Chancellor List of Academic Excellence, 1x IT Award Academic Excellence. Graduated with High Distinction. John Oglethorpe Medal for highest GPA of all IT students graduating.
2006 to 2009	**Bachelor of Life Sciences**, University of Muenster, Muenster, Germany Studied general life sciences with a focus on microbiology. In my final project I worked on EST-based differential gene expression in seagrasses.

Employment

2022 to current	**Research Officer, AI** I lead interpretable AI efforts using genomics and flow cytometry data from WA to assess ocean health.
2021 to 2022	**DECRA Fellow** I used interpretable AI to study mechanisms of gene loss and birth in crops to learn where new genes come from, and how to avoid loss of agronomically important genes. Co-developed a new MSc in Bioinformatics, designed and taught two units.
2017 to 2020	**Forrest Fellow** One of three inaugural Forrest Fellows. Worked on genomics of complex plants with Forrest Foundation support. Taught one bioinformatics unit.
2015 to 2017	**Postdoctoral researcher** Researched genetics of complex plants with a focus on canola and wheat in Edwards lab.

Recent roles

2023	**Founder, WA eDNA meetup** Irregular meetup to connect eDNA practitioners in Western Australia
2022	**Member, Scientific Advisory Panel** Oversee and advise on ongoing industry data collaborations across WA.
2022	**Founding member** Co-founder and lead of teaching.
2021	**Member, Scientific Advisory Panel Machine Learning** Advise on ongoing machine learning projects supported by the ARDC.
2021	**Member, GRDC Machine Learning Technical Consultation Group (ML TCG),** Member of the ML technical consultation group for GRDC-funded ML projects.
2017	**Hacky Hour Founder** Founded a weekly get-together of researchers and staff working with programming and data, doubles as a help-desk for students with programming problems.
2013	**Certified Carpentries Instructor** Certified Software Carpentry and Data Carpentry instructor

Recent presentations

2023	**Perth Machine Learning Group** - Unlocking the power of large language models	
2023	**WA Data Science Innovation Hub (WADSIH)** - My journey towards data skills	
2022	**ResBaz/Perth Machine Learning Group** - Ask us anything - Advice on how to use machine learning to tackle your project	
2022	**Telethon Kids Institute** - Career Pathways Event - how I got here	
2022	**ResBaz** - Bioinformatics showcase	
2021	**CCDM/Curtin University** - Machine learning in bioinformatics – where are we and what's next?	
2021	**Cinvestav/online** - Machine learning in plant breeding and bioinformatics	
2021	**UWA DVCR Forrest Fellow series** - Future-ready crops for a changing climate: the role of bioinformatics	
2021	**Pawsey Supercomputing Centre** - Bioinformatics at scale Q and A	
2021	**ABACBS online seminars** - Interpretable Machine Learning in Bioinformatics	
2020	**GRDC Tech Seminars** - Our machine learning technical stack	
2020	**PAG Conference, San Diego** - Predicting Gene Loss in Plants: Lessons Learned from Laptop-Scale Data	
2019	**Bayliss Seminar Series** - Eukaryotic pangenomics: where we've been, where we're going	
2019	**AGRF Seminar Series** - Assembling complex plant genomes – things I wish someone would have told me earlier	
2019	**PAG Conference, San Diego** - Helping Biologists Make Sense of Plant Variant and Annotation Data	

Academic funding

2020	**Grant: ARC Discovery Early Career Research Award** Awarded DECRA for 2021-2023. Total funding $866k which includes a $419k contribution from UWA.
2020	**Grant: Identifying genetic contributors to canola blackleg resistance in the presence of environmental effects using Machine Learnin** With Prof. Dave Edwards, Prof. Mohammed Bennamoun, Prof. Farid Boussaid, Prof. Jacqueline Batley. Total funding: AUSD 309,524.
2020	**Grant: Machine Learning - Project E: Deep Learning for early detection and classification of crop disease and stress** With Prof. Mohammed Bennamoun, Prof. Farid Boussaid, Prof. Dave Edwards, Dr. Nic Taylor. Total funding: AUSD 344,971.

Relevant publications

1. **Danilevicz** *et al.* (2022). DNABERT-based explainable lncRNA identification in plant genome assemblies. *bioRxiv*, *NA*, NA:NA.
2. **Gill** *et al.* (2022). Machine learning models outperform deep learning models, provide interpretation and facilitate feature selection for soybean trait prediction. *BMC plant biology*, *22*, 1:1–8.
3. **Bayer** *et al.* (2021). Sequencing the USDA core soybean collection reveals gene loss during domestication and breeding. *The Plant Genome (TSI)*, *NA*, NA:1–12.
4. **Bayer** *et al.* (2021). Yield is negatively correlated with nucleotide-binding leucine-rich repeat gene content in soybean. *bioRxiv*, *NA*, NA:NA.
5. **Bayer** *et al.* (2021). Modelling of gene loss propensity in the pangenomes of three Brassica species suggests different mechanisms between polyploids and diploids. *Plant biotechnology journal*, *19*, 12:2488–2500.
6. **Bayer** *et al.* (2020). Plant pan-genomes are the new reference. *Nature plants*, *6*, 8:914–920.
7. **Valliyodan** *et al.* (2019). Construction and comparison of three reference-quality genome assemblies for soybean. *The Plant Journal*, *100*, 5:1066–1082.
8. **Appels** *et al.* (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, *361*, 6403:NA.
9. **Montenegro** *et al.* (2017). The pangenome of hexaploid bread wheat. *The Plant Journal*, *90*, 5:1007–1013.
10. **Golicz** *et al.* (2016). The pangenome of an agronomically important crop plant Brassica oleracea. *Nature Communications*, *7*, 1:1–8.
11. **Chalhoub** *et al.* (2014). Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome. *science*, *345*, 6199:950–953.