Philipp **Bayer**

RESEARCH OFFICER, COMPUTATIONAL BIOLOGY

OceanOmics, Minderoo Foundation

✉ pbayer@minderoo.org | ⬡ 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 am member of two scientific advisory panels, have co-supervised more than twenty PhD and MSc students to completion, published roughly 130 research publications, co-edited a book, and am active in machine learning, eDNA, and open source communities.

At Minderoo OceanOmics I am focusing on pushing the boundary of what is possible using environmental DNA (eDNA) data using genomics and artificial intelligence.

Employment

2022 to current	**Research Officer, Computational Biology**, Minderoo OceanOmics I lead research efforts using interpretable artificial intelligence and genomics to improve how we measure, understand, and protect the ocean.
2021 to 2022	**DECRA Fellow**, UWA I used interpretable AI to study mechanisms of gene loss and birth to learn how to avoid loss of agronomically important genes. I co-developed a new MSc in Bioinformatics at UWA and designed and taught two units.
2017 to 2020	**Forrest Fellow**, UWA I was one of three inaugural Forrest Fellows. I worked on genomics of complex plants with Forrest Foundation support. I taught one bioinformatics unit.
2015 to 2017	**Postdoctoral researcher**, UWA I researched genetics of complex plants with a focus on canola and wheat in the Edwards lab.

Education

2012 to 2016	**PhD**, University of Queensland, Brisbane, Australia Applied bioinformatics PhD in the Edwards group. I developed computational pipeline SkimGBS for higher resolution genotyping while working extensively with industry on plant breeding pipelines (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.

Recent roles

2023	**Founder, WA eDNA meetup** 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.
2013-2023	**Certified Carpentries Instructor** Certified Software Carpentry and Data Carpentry instructor

Recent presentations

I have given a total of 40 public presentations.

2024	**ACS WA conference**	- What we learned using LLMs in environmental DNA
2024	**TEDx Kings Park**	- How we can use social media to protect the oceans
2024	**Curtin University**	- Towards better standards for eDNA taxonomic classification
2023	**WADSIH Data and AI for Business Conference**	- My Journey to a Career in Machine Learning; All My Mistakes You Will Hopefully Avoid
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

Academic funding

I have received a total of $1.5 million in academic research 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.

Selected publications

These represent the most impactful publications out of roughly 130 publications.

1 **Bayer** *et al.* (2024). A comprehensive evaluation of taxonomic classifiers in marine vertebrate eDNA studies. *bioRxiv*, *NA*, NA:2024–02.

2 **Khan** *et al.* (2024). Cicer super-pangenome provides insights into species evolution and agronomic trait loci for crop improvement in chickpea. *Nature Genetics*, *NA*, NA:1–10.

3 **Cundy** *et al.* (2023). Seafood label quality and mislabelling rates hamper consumer choices for sustainability in Australia. *Scientific Reports*, *13*, 1:10146.

4 **Bayer** *et al.* (2022). Not all pathways are the same-unique adaptations to submerged environments emerge from comparative seagrass genomics. *bioRxiv*, *NA*, NA:2022–11.

5 **Danilevicz** *et al.* (2022). DNABERT-based explainable lncRNA identification in plant genome assemblies. *bioRxiv*, *NA*, NA:NA.

6 **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.

7 **Bayer** *et al.* (2021). Sequencing the USDA core soybean collection reveals gene loss during domestication and breeding. *The Plant Genome*, *NA*, NA:1–12.

8 **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.

9 **Appels** *et al.* (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, *361*, 6403:NA.

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.