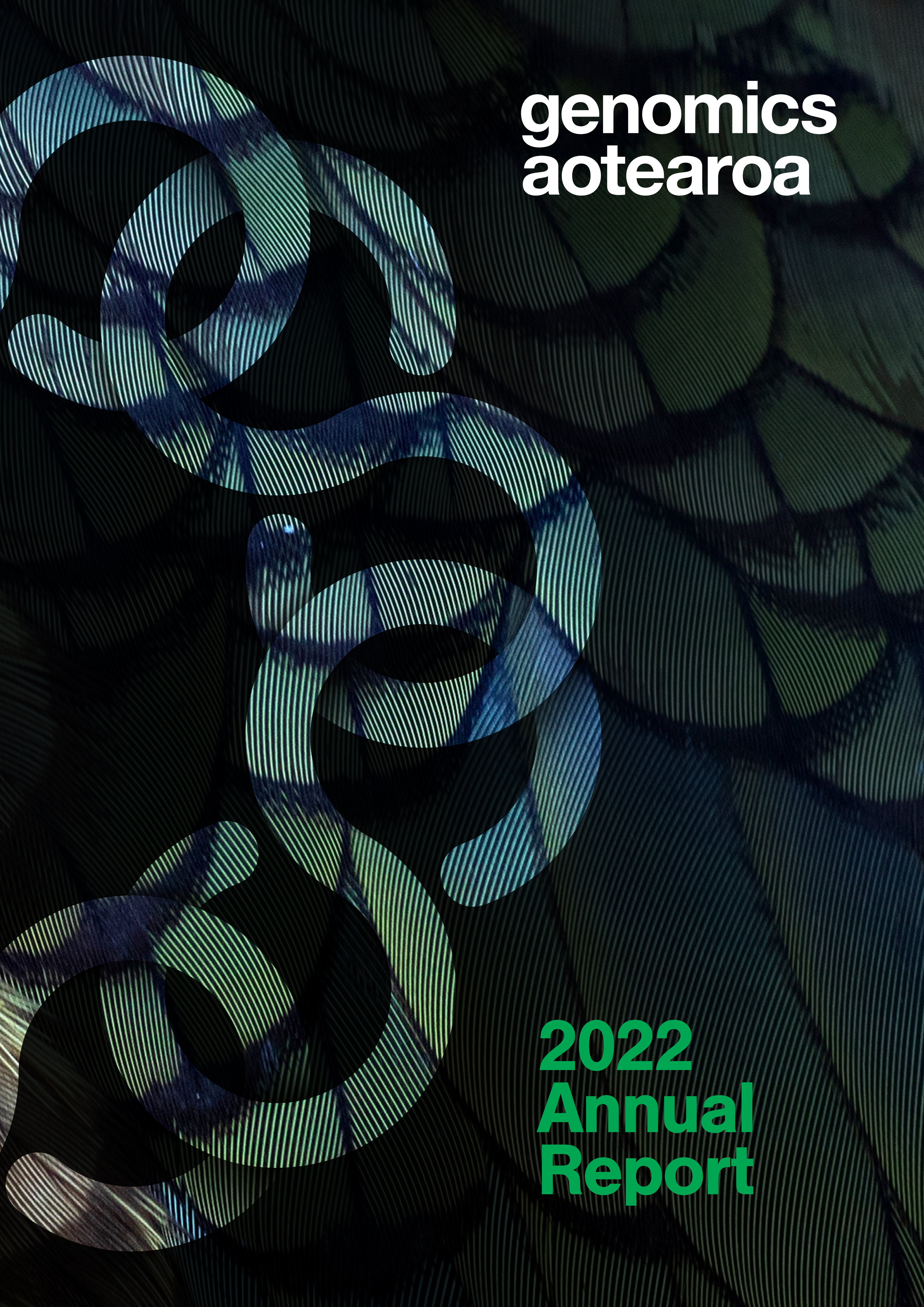


genomics aotearoa

2022 Annual Report

Mā te whakaatu,
kā mōhio

Through discussion,
comes awareness

Mā te mōhio,
ka mārama

Through awareness,
comes understanding

Mā te mārama,
ka mātau

Through understanding,
comes knowledge

Mā te mātau,
ka ora

Through knowledge,
comes wellbeing

Genomics Aotearoa Annual Report

Year 5

1 July 2021 – 30 June 2022

ISSN 2816-0088 (Print)
ISSN 2816-0096 (Online)

"We are cradled in biodiversity, our human world is based on the products of biodiversity and only by understanding and looking after it can we hope to weather this current storm"

— Professor Mark Blaxter

(Tree of Life Programme, Wellcome Sanger Institute), on the impact of climate change

Table of Contents

16. Environment

Report from the Chair


DR WILLIAM ROLLESTON CNZM

Genomics Aotearoa has gone from strength to strength this year. In this report, you will find details of the projects we fund, the people we support, the considerable amount of work undertaken and the outcomes we have generated. This scientific progress is underpinned by changes in the governance of Genomics Aotearoa.

This year we increased Māori representation around the board table. Having key expertise from Te Ao Māori at the highest levels of the organisation, is crucial to building trust with communities and delivering on our commitment to Te Tiriti o Waitangi. I would like to warmly welcome Papatūānuku Nahi, Jan Hania and Gerard O'Regan to the board. While this new board has met virtually a few times, we had a day together in person in June this year, which provided a much-needed opportunity to talk, learn and build relationships.

This year Dr Anna Campbell has announced that she will be stepping down from the board of Genomics Aotearoa, after starting a new business and taking on a leadership role at Otago Innovation Ltd. Genomics Aotearoa benefited hugely from her experience and skills and I wish her all the best for the future.

Genomics Aotearoa staff partnered with Tūhura Otago Museum to develop a genomics exhibition, to be reprised over the summer, and with the potential for touring next year. The exhibition describes aspects of genomics, links these to Aotearoa-specific projects, and shows the benefits of this science in an easy-to-understand way. The Museum project was spearheaded by Operations Manager Jayashree Panjabi, and I congratulate her and her team on this success. This exhibition represents a deliberate and significant move to raise in the public mind the contribution of genomics to all our lives and the significant part New Zealanders have played in that journey.

I would like to thank Peter Dearden and his team for the excellent progress they have achieved over this last difficult year, and the positive outcomes form the mid-term review the scientists who we are fortunate to have in our team and who work tirelessly, patiently, and diligently to progress their work, MBIE for their support and advice, and the board who have provided sage guidance.

It is five years since Genomics Aotearoa was formed, and we can be proud of the achievements to date. The progress to date creates a solid foundation for a strategic review of genomics and bioinformatics in New Zealand, which Genomics Aotearoa, led by the board, will be undertaking over the next 12 months. The review intends to provide the context for future government investment in the critical area of genomics. As this report highlights, genomics and bioinformatics in Aotearoa is thriving and is now a vital part of the local science landscape.

Dr William Rolleston CNZM
Chair of the Genomics Aotearoa Governance Board

Report from the Director

I am pleased and proud to present the Genomics Aotearoa Annual Report for 2021-2022. In the past year our scientists and collaborators have been working hard to deliver the benefits of genomic science to the communities, industries, and agencies we work with.

Genomics is a difficult science. It takes a great deal of creativity to build the resources and tools needed to turn genome data into useful outcomes. Some of the most important aspects of that work are the relationship- and trust-building required to even begin the process. From there the vast quantities of data and the need for better, faster and more effective ways to analyse these data are ever-present challenges.

The effective science communication needed at the end of a project can be even more challenging. Genomics is not a science that can be done hiding in the lab, or behind a computer, generating and analysing data only. Because of this, I am delighted that so many of our projects bring together experts in many areas, especially early-career researchers, which ensures that awesome teams are in place to build trust, build analysis pipelines and finally deliver the benefits of genomics. I hope these teams are an enduring legacy of Genomics Aotearoa's approach.

The key to team building is training, and our training programme, led by Ngoni Faya, has been a real highlight of Genomics Aotearoa. Sadly for us Ngoni is moving to the United States to a new career in human genomics. I would like to thank him for his amazing work and wish him well. Jian Sheng Boey and Tyler McInnes will pick up the reins and I look forward to continued success in this area.

A major highlight this year has been the funding of our first tranche of Indigenous Genomics Platform (IGP) projects. These projects, which aim to build trust in genomics among Māori communities and businesses, are only a start, but I hope we will be able to build on them to produce a cohesive platform. The development of the IGP has been deftly and effectively led by our Vision Mātauranga manager, Tracey Godfery, to whom I am immensely grateful.

PROFESSOR PETER DEARDEN

Another highlight of the year was our mid-term review, which had been delayed by Covid-19, but was held early in 2022. The review report presents a thoughtful and incisive set of recommendations, many of which build into a forward thinking future strategy for Genomics Aotearoa. We are looking forward to the challenges of implementing these recommendations in the next few years.

For the first time in a few years, we met as a community in person in June. A meeting long wished for and successfully achieved, all thanks to our Operations team. Meeting kanohi ki te kanohi (face-to-face) allowed us to begin to refresh our strategy and engage with end-users to build the future of genomics in Aotearoa.

Finally, I write this from a hotel room in Vancouver, the first time in three years I have been overseas for a conference. International links are vital to ensure our work in Aotearoa is at the leading edge of our field and that we are bringing the best science to play for the country. COVID-19 has not gone and is still causing misery and disruption, but the good news is that Aotearoa's genomic science has coped and contributed and will continue to do so into the future.

Professor Peter Dearden
Director of Genomics Aotearoa

Genomics Aotearoa:
A Platform for Collaboration

Since 2017 the Genomics Aotearoa collaborative platform has been leading, facilitating, and promoting genomics research and bioinformatics in Aotearoa. Genomics (the study of the complete genetic information in living things) and bioinformatics (the computational methods and software that enables analysis of genomic data) are our major research disciplines.

Genomics Aotearoa increasingly supports the development of resources by, with and for Māori, and enables greater participation of Māori researchers, communities, businesses, whānau and students in the fields of genomics and bioinformatics.

Genomics Aotearoa has ten partners who represent the genomics expertise of the entire country. The partners include five universities – Auckland, Massey, Otago, Waikato, and Te Herenga Waka-Victoria University of Wellington – and five Crown Research Institutes – AgResearch, Environmental Science and Research (ESR), Manaaki Whenua-Landcare Research, Plant & Food Research and Scion. Thirty-one associate organisations represent researchers and end-users of genomics and bioinformatics across all sectors in Aotearoa that intersect with the life sciences.

With its inclusive nature, the Genomics Aotearoa consortium is in a prime position to lead Aotearoa's genomics research into the future. Genomics Aotearoa and its researchers have established national collaborations with end-users across key themes of health, environment, and primary production. These themes encapsulate the applied life sciences relevant to Aotearoa's economic, environmental, and social wellbeing.

Our end-users include iwi and community groups, government departments and councils, health authorities, hospitals, clinicians, patients and their whānau, breeders and primary producers, and conservation practitioners. The connections between our consortium and end-users ensures that capabilities and discoveries translate to real-world benefits throughout Aotearoa.

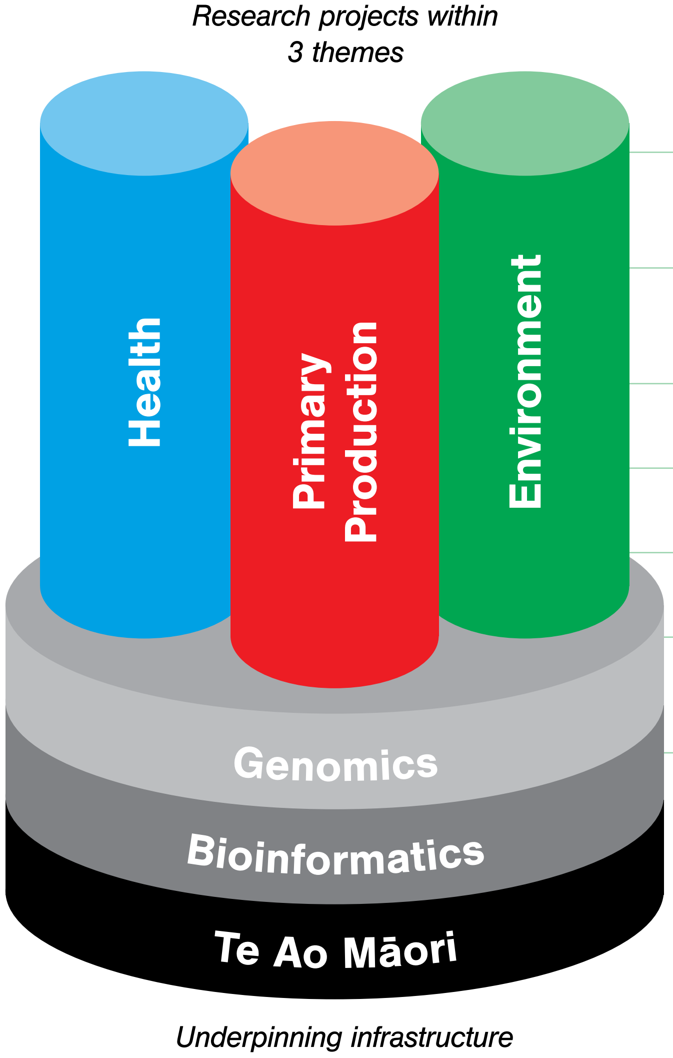
The enabling infrastructure we develop underpins all our research outcomes. Our bioinformatics and genomics capability resides in the more than 130 researchers and other specialists who work with Genomics Aotearoa. Since 2017, the platform has supported 27 postdoctoral fellowships to retain world-class early-career scientists in Aotearoa. Postdoctoral fellows have gone on to become lecturers and lead their own research programmes. Doctoral candidates are involved in many projects, as are undergraduate and graduate students. Extensive genomics and bioinformatics training is available to students, scientists, and clinicians. Worldwide linkages capitalise on, and contribute to, international best practice.

Our cultural infrastructure is essential to our vision to honour our commitments to Te Tiriti o Waitangi. The Aotearoa Genomics Data Repository (AGDR) is essential to enable the appropriate curation and assignation of cultural rights to genomic data. This unique database is increasingly being used as a research data source. The Te Nohonga Kaitiaki guidelines and additional resources being developed represent best-practice for inclusive research.

Partnership with New Zealand eScience Infrastructure (NeSI) enables our research delivery. The AGDR and our bioinformatics platform, both hosted by NeSI, are free for all researchers in Aotearoa to use.

Genomics Aotearoa for 2017-2024 is supported by funding from the Strategic Science Investment Fund, through the Ministry of Business, Innovation and Employment (MBIE). At the beginning of 2022, as planned, MBIE reviewed Genomics Aotearoa's progress to date. Acknowledging our projects' achievements, the reviewers also recognised several aspirations to work towards over coming years.

Strategic needs identified by the review included greater end-user participation in strategy design, planning for the development and permanence of resources such as the AGDR and bioinformatics platforms, more cohesive engagement to develop and promote the capability of Māori researchers, industry representatives and communities, as well as increasing the scope of benefits to include Pacific communities. Fulfilling these strategic needs will ensure our work continues to be relevant to solve real-world problems and achieve long-term social impacts for everyone in Aotearoa.



Pathways to impact

- Increase adoption, impact, and acceptance of genomics and bioinformatics
- Create and indigenous genomics platform including research and data management
- Grow NZ genomics capability through upskilling and capability building
- Grow international connections between individuals and organisations
- Create new national connections between researchers and end users
- Provide an agile, leading-edge collaboration platform of research

Genomics Aotearoa in a Nutshell

TE AO MĀORI

3 Māori governance board members appointed

13 Additions to the Aotearoa Genomics Data Repository
in total – up from nine last year ~**50%** increase

3 Indigenous Genomics Platform projects this year

210 Māori variomes sequenced
(up from **200** last year)

23 Summer Internships for Indigenous Genomics this year

RESEARCH + RESEARCHERS

10 Alliance partners, led by University of Otago

165 Researchers and specialists
34 new this year, from **26** institutions

11 Ongoing projects

31 Associate organisations

3 New projects for next year

TRAINING

Bioinformatics training provided to partner and associate organisation **free of charge**. Courses delivered to date: **52**

394 participants in **20 events** trained this year

Metagenomics summer school attendees to date: **>80**
from **23** research institutes

Trained researchers: **1049**
to date from **21** Universities and Research Institutes

OUTPUTS

80 Peer-reviewed publications to date

864 Citations to date, with **597** within the last 2 years

25 Case studies demonstrating real-world impact

17 Seminars this year

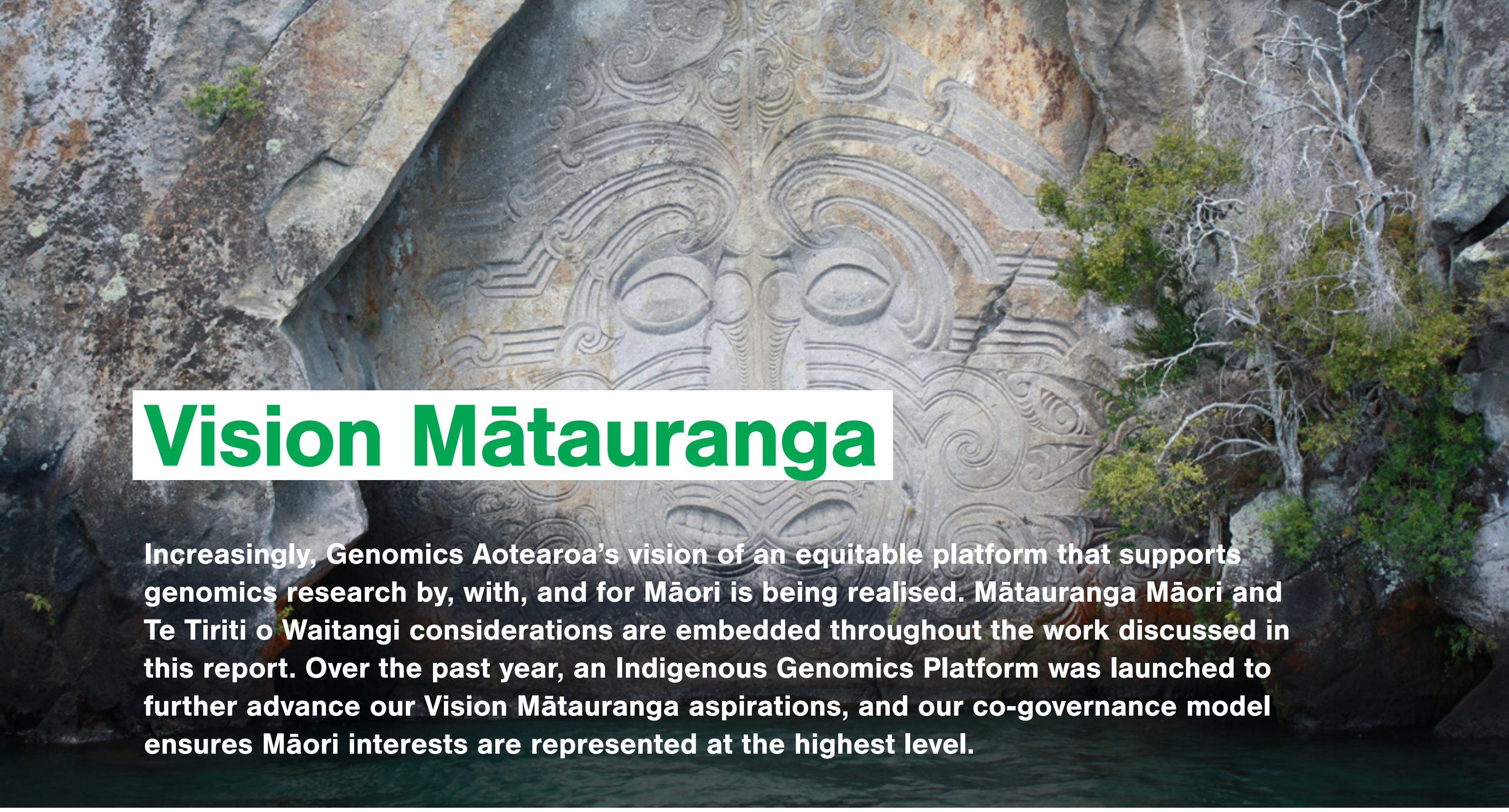
37,852 Visits to the website this year

26 Publications this year

YouTube statistics:
4,963 views
822.2hrs of watch time
76,009 impressions[1]

---

[1] Impressions = times people have been recommended our video or seen it in search



Vision Mātauranga

Increasingly, Genomics Aotearoa's vision of an equitable platform that supports genomics research by, with, and for Māori is being realised. Mātauranga Māori and Te Tiriti o Waitangi considerations are embedded throughout the work discussed in this report. Over the past year, an Indigenous Genomics Platform was launched to further advance our Vision Mātauranga aspirations, and our co-governance model ensures Māori interests are represented at the highest level.

Our co-governance model elevates Māori leadership within Genomics Aotearoa

In our previous report we outlined a co-governance model to increase the representation of Māori partners on the Genomics Aotearoa board. Half the current board are Māori active in their respective iwi and professional sectors – Jan Hania, Papatūānuku Nahi and Dr Gerard O'Regan.

Introducing our new board members

Jan Hania

(Ngati Tuwharetoa, Ngati Raukawa-ki-te-Tonga, Te Ātiawa)

Jan Hania is Environmental Director at NEXT Foundation. He leads and oversees NEXT's environmental portfolio and has taken an integral role setting up the Taranaki Mounga Project, Te Manahuna Aoraki and Predator Free Wellington. Jan is a Trustee on the Cawthron Institute Trust Board and is on the Mana Rangatira of the Bioheritage National Science Challenge programme Te Rakau Taketake. Prior to joining NEXT, Jan spent six years with the Department of Conservation developing large-scale collaborative projects focused on people, biodiversity, and water. His background is in Natural Resources Engineering.

Jan emphasises the benefit of genomics research to Māori, and the necessity for Māori leadership in the field: "Genomics research is a huge opportunity to help solve many of the pervasive challenges we are experiencing, like climate change, biodiversity loss and health. But it is also a potentially significant risk if not used safely and with appropriate involvement, leadership, and determination by iwi Māori," Jan says.

Dr Gerard O'Regan

(Ngāi Tahu)

Gerard O'Regan is our mana whenua board representative. Gerard has worked in heritage management for 30 years. Gerard is Curator Māori and Pouhere Kaupapa Māori at the Otago Museum, and an Honorary Senior Research Fellow in the University of Otago's Archaeology programme. He has cared for Māori collections at the National Museum, provided regional services to community museums and art galleries in the lower South Island, researched bicultural developments in museums, and was a member of the Māori Heritage Council of New Zealand Historic Places. Gerard's MA and PhD at the University of Auckland explored Māori belief of place through the archaeological context of rock art.

Gerard is excited about how genomic research enhances our understanding of the past. "Historically, archaeology gave us tools to discard erroneous, non-sensical and often racist representations of the past. Genomics is unlocking insights into who our ancestors were. Genomics research on animals and plants that our tūpuna lived with is filling in details for what is still an incomplete story of our past interactions and modification of the environment. Treading carefully and sensitively, genomics research is set to write new chapters on our history, both globally and within Aotearoa," says Gerard.

Papatūānuku Nahi

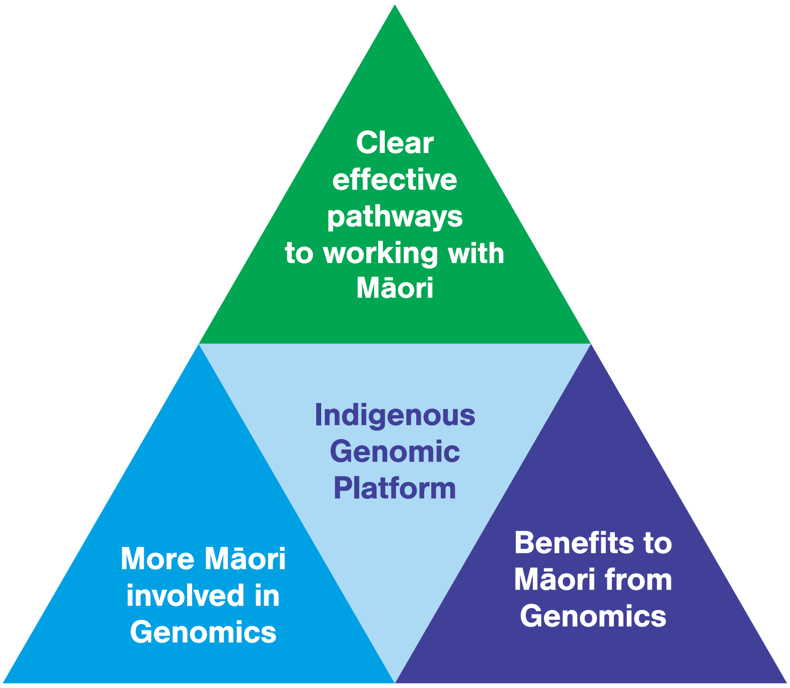
(Ngāpuhi nui tonu)

Papatūānuku (Papa) Nahi is Kai Okohāpai General Manager Tiriti & Equity at Te Hiringa Hauora/Health Promotion Agency. After 20 years in health promotion, she now spends her days on business planning, health strategy development and health service management. Papa is particularly passionate about indigenous innovation and exploring what lies on the horizon. She has been General Manager Māori Public Health at Hapai te Hauora Tapui, and Manager of the Moko Foundation. Papa holds a Master's degree in Health Science (Public Health) with First Class Honours and is a regular presenter to the United Nations indigenous and health committees in Geneva and New York.

"As tangata whenua in Aotearoa, Māori are interested in shaping the landscape, kaupapa and processes of genomics, where there may be impacts on our people, taonga and whenua and to do so in ways that align with mātauranga Māori and our intergenerational focus," says Papa.

An Indigenous Genomics Platform (IGP) as a springboard for Māori-led, Māori-centric research

This year saw the development of our Indigenous Genomics Platform (IGP). The IGP promotes a consolidated approach to research and engagement with Māori and aims to develop clear pathways to increase the participation of Māori in genomics, and ensure Māori receive the benefits of the research.



Clear effective pathways

- Delineate pathways for working with Māori to reflect diverse needs and aspirations
- Cultural, environmental, health, economic

More Māori involved in Genomics

- Capability building at community, student and researcher level
- Education delivery to build trust and increase understanding of genomics

Benefits to Māori

- Grow enterprise
- Improve health and wellbeing at multiple levels
- Knowledge acquisition to enable informed participation in genomics activities

The IGP supports Genomics Aotearoa research and activities to:

- Create an enduring foundation for Māori engagement with genomics that recognises and respects Te Ao Māori values and beliefs.
- Deliver a ground-up approach that provides Māori with a level of trust and understanding to encourage engagement with genomics and genomics researchers.
- Build and promote reciprocal relationships between researchers and participant communities whereby Māori community needs are prioritised.
- Demonstrate Genomics Aotearoa's intent and commitment to fulfil the stated principles, philosophy, and approaches of the Consortium Agreement regarding benefits to Māori.
- Extend across all Genomics Aotearoa activities, irrespective of Māori involvement.

The IGP signals a unique indigenous foundation that encourages respectful and reciprocal engagement with all users of genomics research and infrastructure in Aotearoa. You can read more about the IGP, and the projects funded so far in the Upcoming projects section of this report.

Highlights from our Annual Meeting

The Genomics Aotearoa annual meeting was held June 27-29 at AgResearch Invermay. More than 75 researchers, associates, students, and end-users took part in two days of sharing ideas.

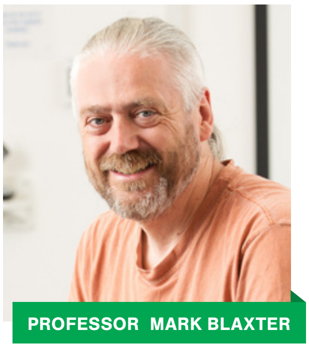
A highlight of the meeting was the plenary talk by Professor Mark Blaxter *Sequence everything: chromosomally complete genomes across the tree of life*. Mark leads the Wellcome Sanger Institute's Tree of Life Programme, which aims to collaboratively sequence the genomes of all species on Earth, under the umbrella of the Earth BioGenome Project. Mark shared the progress, challenges, and interesting insights that the project has revealed.

Later in the meeting, early-career researchers shared their diverse research perspectives. Tom Oosting (Te Herenga Waka-Victoria University of Wellington) joined us remotely from Europe to talk about some of the issues around using a single sequencing platform to assemble genomes. We heard Michael Hoggard (University of Auckland) discuss the influence of environmental viruses on the microbial dynamics of aquatic ecosystems. Joseph Guhlin (University of Otago) talked about the challenges of bioinformatics for small population data and the application of probabilistic modelling. Catherine Collins (University of Otago) talked about her work investigating kiore to understand the movement of people across the Pacific. Christina Straub (ESR) talked about aspects of her *Neisseria* and Group A *Streptococcus* (GAS) research, while Tira Phillipson-Puna (National Hauora Coalition) discussed how the GAS research project engages with its communities.

The afternoon was rounded out by a lively panel discussion with end-user representatives. Panel members included Amanda Black (Tūhoe, Whakatōhea, Whānau-ā-Apanui, Professor at Lincoln University and Director of Bioprotection Aotearoa), Dan Tompkins (Science Director at Predator Free 2050), Darrell Lizamore (Principal Research Scientist at Bragato Research Institute) and Jo Martindale (Molecular section lead at the Wellington Regional Genetics Laboratory). The panel discussed a wide range of topics including the potential of genomics in their industries and where the challenges are. The group emphasised the importance of engaging end-users in project development.

The Genomics Aotearoa/NeSI team held a training day for postdoctoral and other affiliated researchers the day after the meeting. Twenty researchers took part in learning about project and data management on NeSI, use of GitHub, and intermediate-level Shell scripting for bioinformatics projects.

We look forward to another kanohi ki te kanohi meeting next year.



PROFESSOR MARK BLAXTER

Our Partners and Associates

Genomics Aotearoa is an alliance of ten partners:

University of Otago
Te Whare Wānanga o Otāgo
New Zealand

agresearch
āta mātai, mātai whetū

Manaaki Whenua
Landcare Research

The University of Auckland
Te Whare Wānanga o Tāmaki Makaurau
New Zealand

E/S/R
Science for Communities

Massey University
Te Kunenga ki Pūrehuroa
University of New Zealand

Plant & Food™ Research
Rangahau Ahumāra Kai

The University of Waikato
Te Whare Wānanga o Waikato

scion™
FORESTS ▪ PRODUCTS ▪ INNOVATION

Victoria University of Wellington
Te Herenga Waka

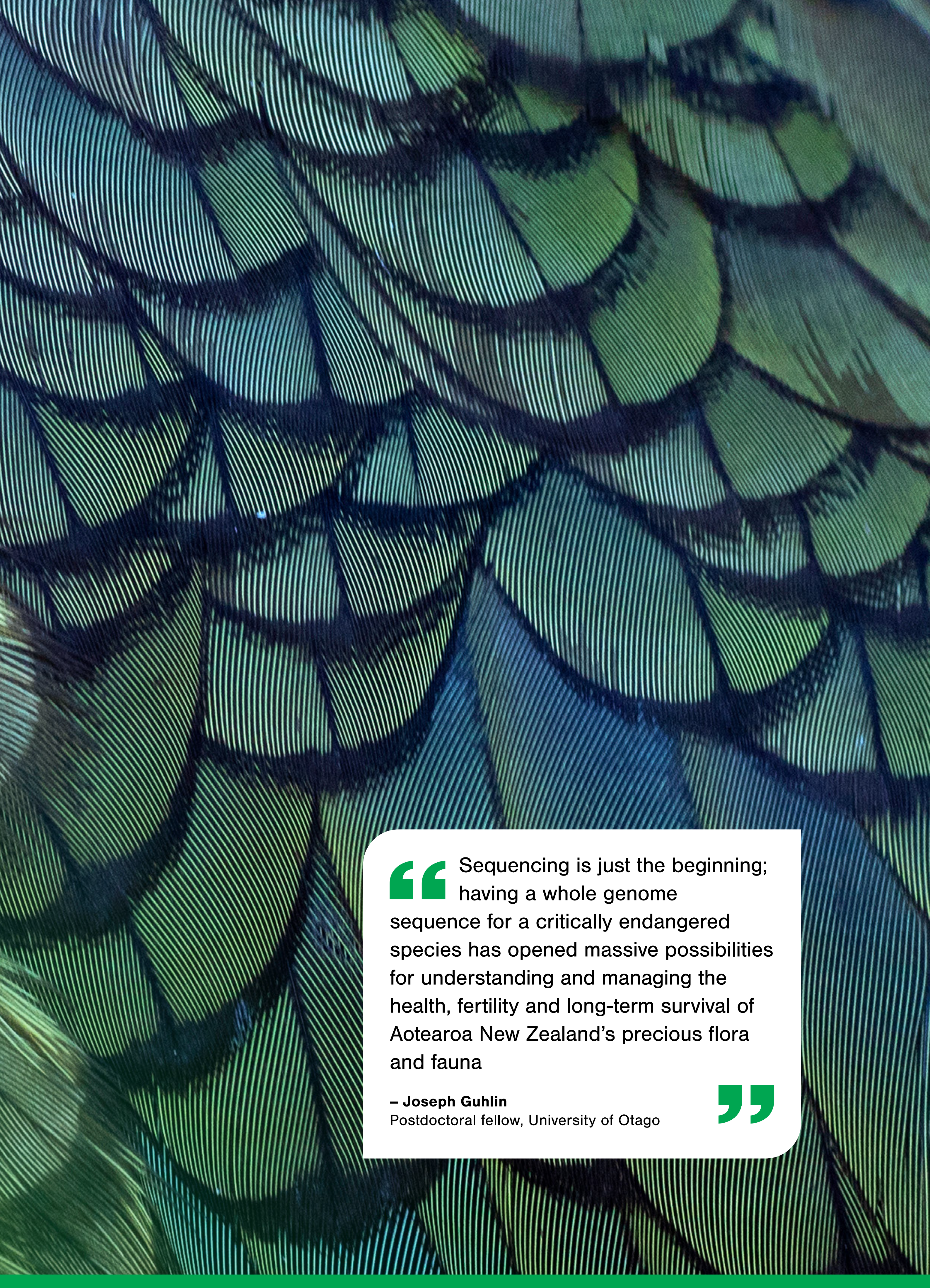
Genomics Aotearoa's 31 associates:

- AbacusBio
- AngusPure
- AsureQuality
- Auckland City Hospital
- Auckland University of Technology
- Beef + Lamb New Zealand Genetics
- Biomatters
- Bioprotection Aotearoa
- Bragato Research Institute
- Brain Research New Zealand
- Cardiac Inherited Disease Group
- Cawthron Institute
- Garvan Institute of Medical Research
- Growing Up in New Zealand
- Hill Laboratories
- LabPLUS
- Livestock Improvement Corporation
- Maurice Wilkins Centre for Molecular Biodiscovery
- Multi-Ethnic New Zealand Study of Acute Coronary Syndromes
- NIWA
- New Zealand eScience Infrastructure
- New Zealand Merino
- New Zealand's Biological Heritage National Science Challenge
- Orion Health
- Pastoral Genomics
- Prevar
- Real Time Genomics
- Royal College of Pathologists of Australasia
- Sophia Genetics
- University of Canterbury
- Waitemata District Health Board

Environment

Environment theme projects use new approaches to unravel some of the most complex genomics problems.

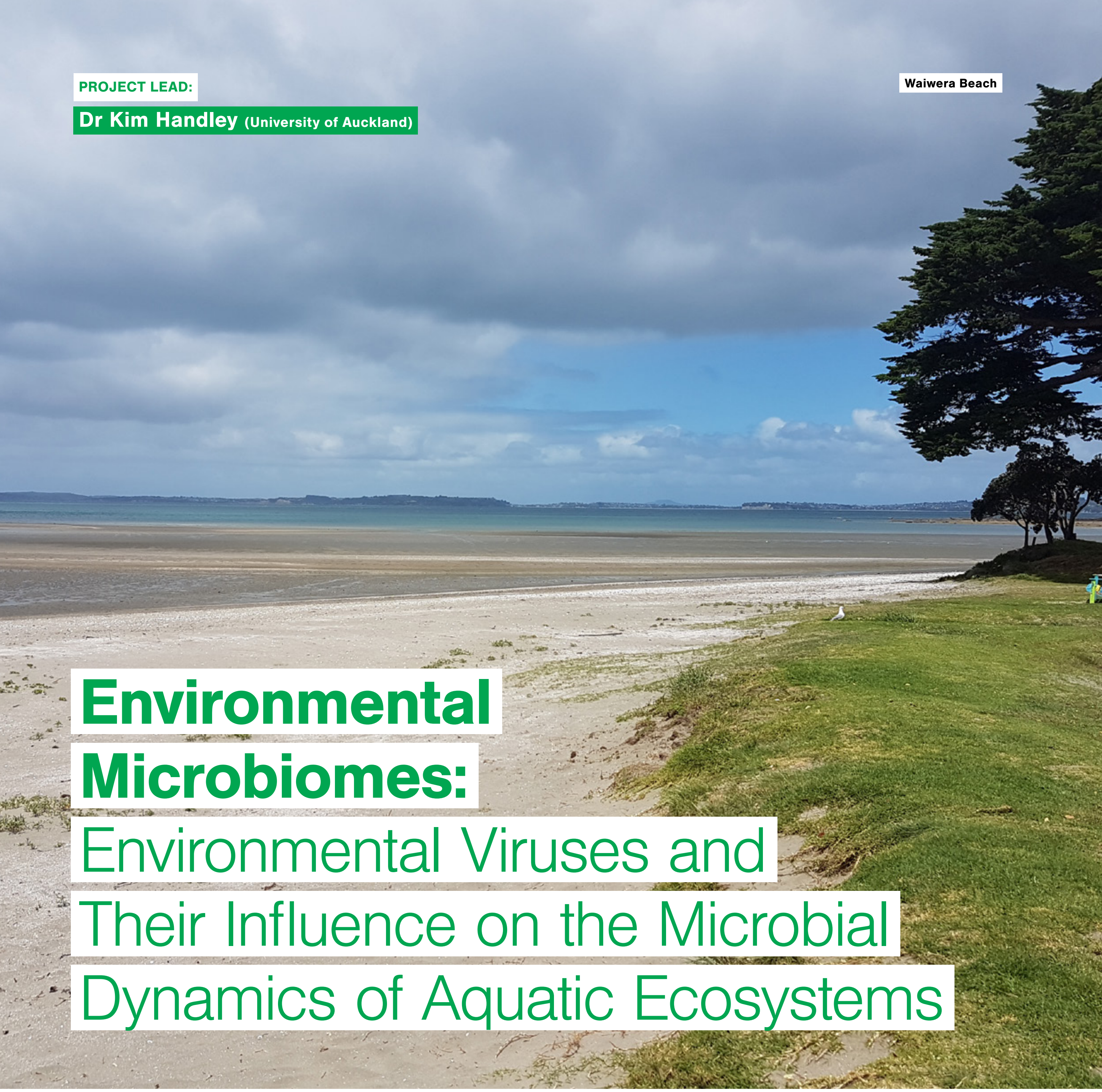
The projects that fall under the Environment theme have sequenced the metagenomes of multiple species in aquatic environments along environmental gradients, large, complex insect and plant genomes, and key invasive species such as wasps and rats. Along the way a significant number of new tools, technologies and pipelines have been explored, refined, and shared among the genomics community of Aotearoa.



> “Sequencing is just the beginning; having a whole genome sequence for a critically endangered species has opened massive possibilities for understanding and managing the health, fertility and long-term survival of Aotearoa New Zealand’s precious flora and fauna”
>
> **– Joseph Guhlin**
> Postdoctoral fellow, University of Otago

PROJECT LEAD:

Dr Kim Handley (University of Auckland)


Waiwera Beach

Environmental Microbiomes:

Environmental Viruses and Their Influence on the Microbial Dynamics of Aquatic Ecosystems

Metagenomics is rapidly transitioning into a mainstream tool in microbial ecology and is now routinely applied to understand communities of prokaryotic organisms (Bacteria and Archaea). The Environmental Microbiomes project extends the scope to include the DNA viruses that infect these prokaryotic organisms

The completed Genomics Aotearoa Environmental Metagenomics project generated massive amounts of data from Waiwera river samples, which is now helping to reconstruct ecosystem composition and function from river-to-sea. These data have been used in several additional studies and projects. A University of Auckland Faculty of Science Fostering Collaboration Award enables Kim Handley and Hwee Sze Tee to work with Cawthron colleagues to analyse the genetics of cyanophycin. The Environmental Metagenomics project had found the problematic bloom-forming cyanobacterium, *Microcoleus*, uses cyanophycin to store nitrogen and carbon.

Three further publications resulted from the project this year.[1] These new opportunities and outputs demonstrate how Genomics Aotearoa's completed projects provide the basis for further novel research.

The new Environmental Microbiomes project is developing best practice guidelines for virome analysis, and an approach for viral genome fragment dereplication in multi-sample environmental datasets. This work capitalises on the virus genomes generated from the Environmental Metagenomics project.

Outputs from the new project align with key objectives for Genomics Aotearoa capability building and advanced genomics research, through the contribution of new step-by-step workflows for metagenomics and hands-on-training, and research at the leading edge of microbial ecology.

The project team have completed a viral processing pipeline for the reconstruction, classification, and analysis of DNA viruses from metagenomic data. The methods have been applied to estuarine and groundwater datasets to investigate the understudied viral ecology of these systems.

Another avenue of research is resolving strain heterogeneity in *Aquiluna* bacteria, the most abundant bacterium identified in the Waiwera estuary. The bacterium has a broad distribution and tolerance to salinity. However, this apparent cosmopolitanism could result from a mixture of strains, which may be the reason *Aquiluna* genomes from this site assemble poorly. As well as improving the genome assemblies for *Aquiluna*, the project will experimentally test salt tolerance among the different strains to validate metagenome-based predictions.

The Environmental Microbiomes project has already resulted in several new tools in development or completed for metagenomic analysis. These tools include a DNA viral processing pipeline (for whole genome sequencing), an RNA viral processing pipeline (for whole transcriptome sequencing) and an Oxford Nanopore processing pipeline for sequence data from prokaryote cultures. All these tools contribute to the Genomics Aotearoa bioinformatics platform and upskilling of researchers through the Metagenomics Summer School.

The Environmental Microbiomes project extends our understanding of the dynamics of the world of microbial communities hidden to us prior to the advent of metagenomics. As estuarine and groundwater resource quality is increasingly at threat, such understanding is increasingly important.

Waiwera Stream

Waiwera River

---

[1] Tee HS, Wood SA, Bouma-Gregson K, Lear G, Handley KM. 2021. Genome streamlining, plasticity and metabolic versatility distinguish co-occurring toxic and non-toxic cyanobacterial strains of *Microcoleus*. mBio 12: e02235-21. doi.org/10.1128/mBio.02235-21

Tee HS, Waite DW, Lear G, Handley KM. 2021. Microbial river-to-sea continuum: gradients in benthic and planktonic diversity, osmoregulation and nutrient cycling. Microbiome 9:190. doi.org/10.1186/s40168-021-01145-3

Mosley OE, Gios E, Weaver L, Close M, Daughney C, van der Raaij R, Martindale H, Handley KM. 2022. Metabolic diversity and aero-tolerance in anammox bacteria from geochemically distinct aquifers. mSystems 7:e01255-21. doi.org/10.1128/msystems.01255-21

CASE STUDY:

Using viral metagenomics to better resolve the function and health of our waterways

Direct sampling of microorganism DNA from the environment is a powerful way of determining the biological diversity and metabolic potential of microbes in an ecosystem, with implications for our understanding of the functioning and health of our waterways.

Environmental DNA (eDNA) data are being used to research the genomes of simple viruses to determine their function in aquatic health, to understand host interaction between viruses and the host bacteria they infect, and the influence viruses have on the aquatic ecosystem.

Postdoctoral fellow Michael Hoggard wanted to understand what viruses are found in the Waiwera River, their interactions with bacteria, and how this affects the ecosystem dynamics, particularly their role in nutrient cycling.

“We can’t manage waterways without this core knowledge – genomics has given us the tools to understand basic biology at a fundamental level and build from there. This is becoming the foundation for understanding and maintaining water quality,” says Michael.

“It’s an on-going story but we’re continuing to fill in gaps, gaining a good understanding of the roles and interactions among bacteria, microorganisms and viruses and their nutrient needs in sediment as well as water.”

This study also provides new insight into virus and host diversity across a salinity gradient from freshwater to marine, as well as biological responses to changes in this highly dynamic environment.

Michael’s analysis recovered more than 30,000 putative viral sequences from the estuarine dataset alone, including 352 complete or near-complete viral genomes spanning five viral families. The research is so groundbreaking that most viral genomes recovered were not able to be assigned to a known taxon due to their novelty. Such novel results highlight the power of genomics to uncover hidden biological diversity.

Matching the recovered viruses to specific bacterial hosts is a challenge in a dataset of environmental samples. The project’s analyses have been able to identify candidate bacterial hosts for the recovered viruses from within a set of bacterial genomes reconstructed from the same dataset.

Analysis of estuarine viral gene expression indicates the viruses modify or enhance a range of the host’s functions, including photosynthesis, biosynthesis, DNA repair and evasion of host defences. Additional analysis will further explore evidence of direct interactions between viruses and bacteria, including bacterial defence mechanisms, the potential effect of viral infection on overall microbial community structure and functioning, and viral manipulation of host metabolism.

Findings to date suggest that viruses can significantly manipulate microbial communities and biogeochemistry in the terrestrial subsurface.

Our new understanding at both the microbial and virus level enables a more comprehensive overview of the Waiwera Estuary composition and functioning.

Very importantly, the research is also providing a model of genomics-based tools that can be applied to understand microbial interactions in other waterways and coastal environments, for use by other researchers, with potential benefit for end-users.

Understanding the basic biology is a fundamental step in long-term planning for managing our waterways. This work has massive potential for improving Aotearoa’s environment.

PROJECT LEADS:

Prof Thomas Buckley (Manaaki Whenua – Landcare Research)

Dr Shannon Clarke (AgResearch)

Dr David Chagné (Plant & Food Research)

Dr Anna Santure (University of Auckland)

Poor Knights Giant Wētā (Deinacrida fallai) – *Credit:* Dylan Van Winkel (CC BY-SA 4.0)

High Quality Genomes and Population Genomics for the Environment

High quality genomes are increasingly being used as tools to help threatened species population recovery. Potential adaptive resilience to environmental instability, such as climate change, requires knowledge of the genomes of entire populations of a species (the pangenome), and their variation (the variome). Increasingly, population genomics is within reach of conservation managers. Honouring Te Tiriti o Waitangi obligations is critical to the success of research on taonga and conservation outcomes.

The High Quality Genomes and Population Genomics project includes sub-projects within both the Primary Production and Environment themes. Additional information on the development of tools and their application can be found in the Primary Production theme, and in the Bioinformatics infrastructure sections. Most of the 43 researchers in the HQG+PG project are associated with the Environment theme.

New technology has been a game-changer for the project. The HQG+PG project has been an early adopter of long read high fidelity Pacbio technology. Using Pacbio has dramatically improved genome assembly, removing the need for time-consuming error correction steps required previously. The project team has developed new genome assemblies for several species, and integrated Pacbio into an optimised bioinformatics workflow. The time to produce a complete, accurate genome assembly has been reduced from months of calculation to a few hours. Furthermore, the technology enables the team to tackle complex genomes more easily, such as the very large genomes of native *Leiopelma* (Hochstetter's frogs), giant wētā and stick insects. The Hochstetter's frog genome (assembled by Manaaki Whenua in collaboration with Ngāti Porou) is the largest and most complex assembled to date by the project team (8 Gigabases).

In previous years the project generated complete assemblies for bilberry, rewarewa, stick insect, and hihi, and this year the project completed a phased chromosome length assembly for karaka. Existing assemblies provide a wealth of raw material for additional research.

Sequencing and analysis for many other species are in progress or nearing completion, including huhu, rātā moehau (an extremely rare white-flowered rātā), kuaka (Whenua Hou Diving Petrel), *Lepidium* species (a group of rare plants), toroa (Antipodean albatross), kiore and *Peripatus* (velvet worms, which are evolutionarily fascinating and not worms at all).

Māori groups are actively involved throughout the research programme

The work on taonga would not be possible without the involvement of mana whenua. Māori kaitiaki interests, tikanga and Vision Mātauranga are incorporated throughout the execution of projects.

Communities and kaitiaki are involved in the approval and collection of all animal and plant samples for a given subproject, as well as discussions on project progress. Communities and kaitiaki involved[2] include representatives from Te Awahohonu Forest Trust (huhu), Ngāti Kuri (rata moehau), Te Rūnanga o Ngāi Tahu (kuaka, *Lepidium sisymbrioides*, takahe [with Tāne Davis] and toroa), Rangitaane o Manawatu (see also the Upcoming projects section) and Te Rūnaka ō Ōtakou (Ngāi Tahu) (*Peripatus*), Hokotehi Moriori Trust (kiore), Ngāti Rangi (pekapeka [with Karen Mitchell of the Ngā Waihua o Paerangi Trust]) and Ngāti Porou (Hochstetter's frog). Connections with iwi throughout Aotearoa for kanakana/piharau (lamprey) sequencing have been facilitated by Dr

Leiopelma hochstetteri – *Credit: Neil Birrell (CC BY 4.0)*

---

[2] More groups are discussed under the Primary Production theme.

Manuka Flowers and Native Bee – *Credit: Avenue* (CC BY-SA 3.0)

Jane Kitson (Ngāi Tahu ki Murihiku [Ōraka-Aparima, Waihopai and Awarua Rūnanga]). This long list demonstrates the commitment of the Genomics Aotearoa's community to upholding commitments to Te Tiriti o Waitangi, the goals of Vision Mātauranga and the broader social license needed for successful research outcomes in Aotearoa.

Working in a way that respects tikanga requires commitment to open sharing of the research approach, and an acceptance that there is a right of participants to decline. Extensive consultation was needed for genomics research on Hector's and Maui dolphin, as these species are distributed over much of Aotearoa's coastal waters. Researcher Alana Alexander consulted with representatives from 39 rūnanga and other iwi/hapū representatives[3] to obtain consent for generation of genomics data. One rūnanga declined consent for dolphins from their takiwā, so samples from their rohe were not included. However, Alana is very grateful to the rūnanga for their time and expertise in offering their whakaaro, and she continues to update them on the project so they can monitor her intentions and engagement in this space.

Increasing the involvement of the community early in the development of projects, or even before the projects are decided upon, is critical not only to project success, but also to the vision of Genomics Aotearoa. Two Māori emerging leaders in conservation genomics, Alana Alexander and Catherine Collins, who lead the Ruatau project, have recently been appointed as lecturers at the University of Otago. For Catherine and Alana, community-embedded projects are vital. Alana has incorporated this approach into her teaching, so that future researchers will also be inspired to encourage community-led projects with researchers embedded within the community.

Peripatoides novaezealandiae – *Credit: Christopher Stephens (CC BY-SA 4.0)*

---

[3] Awarua Rūnanga, Kāti Huirapa Rūnaka ki Puketeraki, Ngāi Tai ki Tāmaki, Ngāti Apa ki te Rā Tō, Ngāti Kōata, Ngāti Kuia, Ngāti Maru (Hauraki), Ngāti Rārua, Ngāti Tama ki Te Tau Ihu, Ngāti Tamaoho, Ngāti Tamaterā, Ngāti te Ata, Ngāti Toa Rangatira, Ngāti Whātua, Ōnuku Rūnanga, Ōraka-Aparima Rūnaka, Rangitāne o Wairau, Te Ākitai Waiohua, Te Atiawa o Te Waka-a-Maui, Te Iwi o Te Roroa, Te Kāhui o Taranaki Trust, Te Kawerau a Maki, Te Ngāi Tūāhuriri Rūnanga, Te Rūnanga o Arowhenua, Te Rūnanga o Hokonui, Te Rūnanga o Kaikōura, Te Rūnanga o Koukourārata, Te Rūnanga o Makaawhio, Te Rūnanga o Moeraki, Te Rūnanga o Ngāi Tahu, Te Rūnanga o Ngāti Waewae, Te Rūnanga o Ōtākou, Te Rūnanga o Waihao, Te Taumutu Rūnanga, Te Uri o Hau, Waihōpai Rūnaka, Waikato-Tainui, Wairewa Rūnanga and Te Rūnanga o Rāpaki

Domestic and international collaborations continue to support project success

Across the board, collaborations continue to be important for the success of the project.

Project team members continue to work alongside Predator Free 2050 for sequencing of key pest species. Genomes for ship rat and possum were completed in 2021 and are nearing publication. The genome assembly of kiore (in collaboration with Hokotehi Moriori Trust) is also nearing completion. Further work is planned in a new project with Predator Free 2050 to develop genomic datasets that allow an understanding of gene flow in Norway and ship rat populations. This will help determine how best to introduce single-sex offspring suppression (SSOS) rats for rat population control and eradication.

Collaboration with the Department of Conservation, together with iwi consent and involvement, is vital to the sequencing of takahē, Maui and Hector's dolphins, kākāpō, and pekapeka.

Many other entities have also contributed to the HQG+PG sub-projects.

Takahē genome sequencing is undertaken in partnership with the International Vertebrate Genome Project and B10K consortium. These partners have increased the international profile of this work, and the team have gained early access to state-of-the-art sequencing, assembly and annotation pipelines and approaches. The takahē sub-project is also supported in part by Revive & Restore Wild Genomes.

Hector's and Maui dolphins are one of the species selected for Oxford Nanopore's ORG.one initiative, a pilot-stage project to support faster, more localised sequencing of critically endangered species. This collaboration is providing MinION flow cells at no cost for sequencing, and international exposure/media coverage of the Hector's and Maui project.

Analysis of populations of swamp maire from around Aotearoa is being carried out using the tools developed in the HQG+PG project, in collaboration with the Bioheritage National Science Challenge programme Te Rakau Taketake. The objective of this research is to save swamp maire from the threat of Myrtle rust disease.

The hihi genomics sub-project, led by Anna Santure, is aligned to the large MBIE Data Science programme *Beyond Prediction: Explanatory and transparent data science,* with hihi genome sequencing supported by Auckland Genomics, University of Auckland.

The kākāpō population genomics sub-project has been carried out with the Vertebrate Genome project and a large set of international and domestics collaborators make up the kākāpō 125 genome consortium.

The kākāpō *Aspergillus* sequencing project involves collaboration with international groups including Jason Stajich (University of California, Riverside) and Mathew Fisher (Imperial College London). This work also involves a large group of vets and microbiologists in the Kākāpō Aspergillosis Consortium. *Aspergillus* is a major threat to kākāpō, and often causes severe respiratory illness, which can be fatal.

The Department of Conservation is funding a genomic assessment of the Antipodean and Gibson's albatrosses. This work will resolve taxonomic uncertainty to aid in the cross-border conservation of these threatened endemic species. The partnership with DOC will allow this research to reach those involved in management of the species, as well as industries such as fisheries that interact with the birds.

Kiore *(Rattus exulans)* – *Credit:* Forest and Kim Starr (CC BY 3.0 US)

Metrosideros bartlettii – *Credit: Bodleye1 (CC BY-SA 4.0)*

The HQG+PG project has led to a wealth of new opportunities

The capabilities in genome assembly developed in this project support a Royal Society Te Apārangi Marsden Fund award to Thomas Buckley (Manaaki Whenua) *Identifying the genomic underpinnings of successful asexuality*. The Marsden-funded work will include assembly of several stick insect genomes using pipelines developed in the HQG+PG project.

Genome graph assembly capability led to a MBIE Endeavour Fund (Smart Ideas) award to Tammy Steeves (University of Canterbury) *Building an integrative genomics framework to mitigate maladaptive reproductive traits in endangered species for kākāpō*, which will use pipelines from the HQG+PG project to develop genome graph assemblies for the species.

Genomics expertise developed at University of Auckland through the HQP+PG project has led to a new project for Anna Santure to assess population structure and connectivity of the threatened New Zealand Storm Petrel. This partnership with the New Zealand Storm Petrel working group and Northern New Zealand Seabird Trust has been funded by a BirdsNZ research grant. BirdsNZ also supported a project to enable PhD candidate Marissa Le Lec to develop new Nanopore based tools for rapid determination of sex and parentage of kākāpō.

The breadth of research effort in, and outstanding science resulting from, the HQG+PG project is demonstrated by the number of resulting publications, including 15 related to the Environment theme alone.[4] The exemplary science within the project is embodied by Jana Wold's perspective article *Expanding the conservation genomics toolbox: Incorporating structural variants to enhance genomic studies for species of conservation concern* in *Molecular Ecology*.[5] This article earned Jana recognition as second runner-up to the 2022 Harry Smith Prize. She was invited to join the Junior Editorial Board of *Molecular Ecology* and *Molecular Ecology Resources,* both top ranked journals. The article also involved GA-affiliated researchers Anna Santure, Tammy Steeves, Joseph Guhlin and Marissa Le Lec.

---

[4] Wu C et al. 2021. Divergent gene expression following duplication of meiotic genes in the stick insect *Clitarchus hookeri*. Genome Biology and Evolution, 13: evab060. doi.org/10.1093/gbe/evab060

Cheng H et al. 2022. Haplotype-resolved assembly of diploid genomes without parental data. Nature Biotechnology. doi.org/10.1038/s41587-022-01261-x

Morrissey KA et al. 2022. Comparison of reptilian genomes reveals deletions associated with the natural loss of γδ T cells in squamates. The Journal of Immunology 208: 1960-1967. doi.org/10.4049/jimmunol.2101158

Dussex N et al. 2021. Population genomics of the critically endangered kākāpō. Cell Genomics 1: 100002. doi.org/10.1016/j.xgen.2021.100002

Martini d et el. 2021. Evolution of the "world's only alpine parrot": Genomic adaptation or phenotypic plasticity, behaviour and ecology? Molecular Ecology 30, 6370-6386. doi.org/10.1111/mec.15978

Duntsch L et el. 2022. Who are you? A framework to identify and report genetic sample mix-ups. Molecular Ecology Resources 22: 1855-1867. doi.org/10.1111/1755-0998.13575

Lee KD et el. 2022. The design and application of a 50K SNP chip for a threatened Aotearoa New Zealand passerine, the hihi. Molecular Ecology Resources 22: 415-429. doi.org/10.1111/1755-0998.13480

Galla SJ et el. 2022. The relevance of pedigrees in the conservation genomics era. Molecular Ecology 31: 41-54. doi.org/10.1111/mec.16192

McCulloch GA et el. 2021. Genomic signatures of parallel alpine adaptation in recently-evolved flightless insects. Molecular Ecology 30:24, 6677-6686 doi.org/10.1111/mec.16204

McCulloch GA et el. 2021. Genomics reveals widespread ecological speciation in flightless insects. Systemic Biology doi.org/10.1093/sysbio/syaa094

Leask MP et el. 2021. Evolution and genomic organization of the insect sHSP gene cluster and coordinate regulation in phenotypic plasticity. BMC Ecology and Evolution 21, 154 doi.org/10.1186/s12862-021-01885-8

Magid M et el. 2022. Leveraging an existing whole genome resequencing population dataset to characterize toll-like receptor gene diversity in a threatened bird. Molecular Ecology Resources. Accepted Article. doi.org/10.1111/1755-0998.13656

Forsdick N et el. 2022. Current applications and future promise of genetic/genomic data for conservation in an Aotearoa New Zealand context. Science for Conservation. In Press. doi.org/10.32942/osf.io/7bnjw

Hauser S et el. 2022. Comparing genome-based estimates of relatedness for use in pedigree-based conservation management. Molecular Ecology Resources. Early View. doi.org/10.1111/1755-0998.13630

[5] Wold J et el. 2021. Expanding the conservation genomics toolbox: incorporating structural variants to enhance genomic studies for species of conservation concern. Molecular Ecology 30: 5949-5965. doi.org/10.1111/mec.16141

Zephyr-A2-2022 (left) and Pearl-A2-2022 (right), demonstrate how quickly kākāpō chicks grow up. Zephyr is two days old (the youngest chick of the season), and Pearl is 38 days old (the second oldest chick of the season). – *Credit:* **Lydia Uddstrom**

CASE STUDY:

Kākāpō recovery – understanding genetic characteristics critical to species survival

The kākāpō is a unique, critically endangered parrot living in the wild only on predator-free islands in Aotearoa. Intensive conservation management saw the population recover from a low of 51 individuals in 1995 to about 150 adults in 2019.

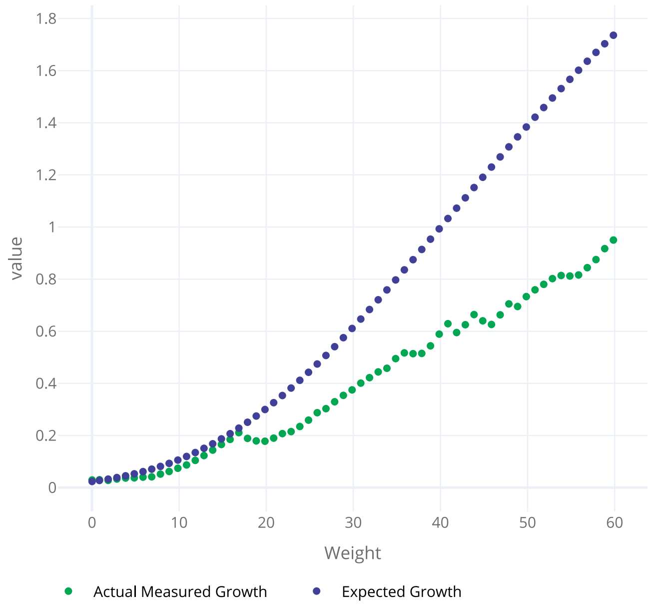
Sequencing of the genomes of every kākāpō was completed in late 2018 through the kākāpō 125+ project, established by the Department of Conservation with funding organised by the Genetic Rescue Foundation.

However, genomics technology has advanced in those few short years, and in 2022 the Genomics Aotearoa High Quality Genomes and Population Genomics (HQG+PG) project produced a high-quality genome for kākāpō. Joseph Guhlin and his team from around the country developed state-of-the-art methods to detect variation among individuals to a quality that meets the rigorous standards of human genome mapping.

With the kākāpō genomes now totalling 169, the new high-quality variant dataset offers researchers the tools needed to identify specific genetic characteristics that are critical to species survival.

The variant calling pipeline developed for kākāpō exceeded the expected quality for similar populations and is being further developed into reusable code for all researchers to use. The high-quality variant dataset is available for kākāpō researchers to use through the Aotearoa Genomic Data Repository.

Morehu Growth Actual vs Expected




Kākāpō – *Credit:* Andrew Digby

The study was able to link a range of variants to phenotypic differences among individuals. Because of this variant-calling work a wealth of information is now available on genetic variations affecting growth, fertility, embryo survival, and clutch size. This new knowledge has:

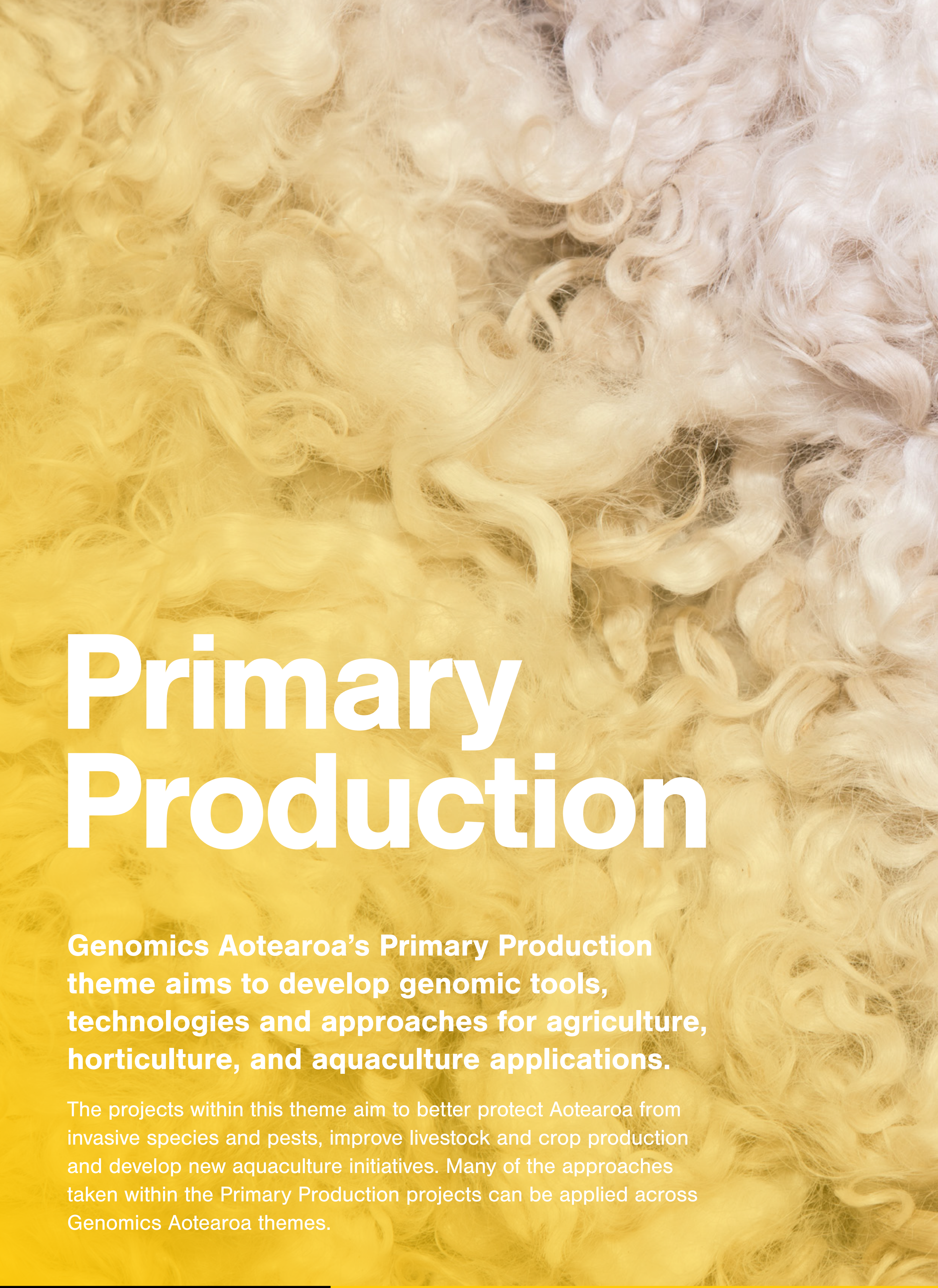
- Revealed genetic contributions to egg shape, which provides a baseline test to validate that methods are appropriate for the small sample size and agrees with findings for other bird species.
- Identified variants of interest to make it easier to identify chicks most at-risk from other diseases and conditions, such as aspergillosis susceptibility and egg fertility.
- Helped understand how genetics contribute to early growth. This understanding can lead to more targeted monitoring, the potential for intervention if a chick is not thriving, and prioritising chicks for special care.

The kākāpō high-quality genetic variant resource provides a much-needed foundation for understanding and translating genomics for future generations of these precious birds. The depth of genetic information now available enables conservation biologists to use genomics as a practical tool to estimate population structure and genetic diversity, as well as respond to problems like infertility and diseases.

“This new tool assists conservationists to identify potential illness at an earlier stage if a chick deviates from their expected growth rate.” says Lydia Uddstrom, Veterinary Advisor, Kākāpō at the Department of Conservation.

Adapting and developing these new genomic approaches has promise for managing other threatened species in more powerful ways than ever before.


Sirocco – *Credit:* Jake Osborne



Primary Production

Genomics Aotearoa's Primary Production theme aims to develop genomic tools, technologies and approaches for agriculture, horticulture, and aquaculture applications.

The projects within this theme aim to better protect Aotearoa from invasive species and pests, improve livestock and crop production and develop new aquaculture initiatives. Many of the approaches taken within the Primary Production projects can be applied across Genomics Aotearoa themes.

> It’s vital we build our genomic capability for researchers and breeders so they can keep up in an ever-changing world, to retain and improve consumer experiences in Aotearoa New Zealand’s meat and wool industries
>
> – **Shannon Clarke (Ngāi Tahu)**
> Animal Genomics Science Team Leader,
> AgResearch

PROJECT LEADS:

Dr Manpreet Dhami (Manaaki Whenua – Landcare Research)

Dr Angela McGaughran (University of Waikato)

Invasomics for Better Biosecurity and Invasive Species Management

Credit: Davide Bonora/Shutterstock

Invasive species threaten native biodiversity, primary industries, and health. Ever-increasing international trade, mobility, and the threat of climate change make Aotearoa highly vulnerable to newly arriving pests and diseases. Our unique ecosystems add a layer of complexity to our responses to these threats. Innovative and customised solutions that minimise the risk of new pests arriving are therefore urgently needed by policymakers and practitioners.

Aotearoa continually needs tools to better prevent, target, and manage biological invasions. New this year, the Invasomics project seeks to identify and validate measurable genetic characteristics associated with inherent invasiveness. High-risk species that could significantly impact primary production if they arrive in Aotearoa have been chosen as models for the work and include Brown marmorated stink bug (BMSB; *Halyomorpha halys*) and Queensland fruit fly (*Bactrocera tryoni*). These species will be used to identify 'omic[1] signatures that underlie successful biological invasion, and these signatures will be used to predict the invasive potential of the species.

---

[1] 'Omics is the collective name for the disciplines that characterise and quantify pools of biological molecules that translate into the structure, function, and dynamics of organisms. Examples include genomics, proteomics, metabolomics, metagenomics, phenomics and transcriptomics.

The team intends to produce an effective bioinformatics pipeline for others to be able to generate high-quality 'omic data for biosecurity risk species, and a validated, generalisable model framework as a tool to assist with prioritisation of potential invasive species. Working collaboratively, end-users such as the Ministry of Primary Industries (MPI) will be able to incorporate this tool into their current risk categorisation pipeline.

So far, the team has used population genomic techniques to show that the BMSB has a complex invasion history, with multiple introductions across the globe out of China. Invasive populations show a high degree of shared ancestry and contemporary sharing of genetic material, which is a challenge for accurate assignment of new BMSB incursions to their source populations. This challenge is likely to be broadly applicable to other species where there is a high degree of admixture among new populations. The team are currently working on publication of these findings.

The team has several initiatives to share their research outcomes more broadly. The development of an Invasomics Hub (invasomics.com) provides a focal point for biosecurity genomics researchers and end-users. And the team have shared their research methodologies on Github, with a step-by-step tutorial for population genomic analysis of BMSB (github.com/Elahep/BMSB-popgenomics) and a repository to derive an invasiveness score (github.com/AlaMetTyr/Invasivenesstraits).

CASE STUDY:

Engaging with end-users of invasomics to prevent a stinky problem

Brown marmorated stink bug (BMSB; *Halyomorpha halys*) is an insect pest that could decimate our fruit and vegetable industries if it gets here. Alarmingly, during autumn and winter, thousands of these bugs can enter houses to shelter from the cold. And when they are disturbed, the bugs give off a foul-smelling liquid that can make a home difficult to live in.

So, what has Genomics Aotearoa been doing to help stop the brown marmorated stinking bug and other nasty pests at the border?

Manpreet Dhami and Angela McGaughran have been working in partnership with MPI on innovative new solutions to help predict the invasive potential of pests.

This engagement with biosecurity researchers and practitioners has been essential to get the Invasomics project off the ground. Project leaders held workshops in 2022 with MPI representatives, alongside a well-attended Biosecurity Bonanza webinar.

Manpreet found the early involvement of MPI to be essential, “The feedback from MPI provided critical direction to focus our initial efforts on BMSB, and it has strengthened our relationship.”

When a new species is detected post-border, Aotearoa currently has a low capacity for generating a rapid genome-informed management response. The Invasomics project is geared towards resolving this challenge, with the aim of better preventing, targeting, and managing biological invasions.

As Angela says, “We hope that advanced genomic techniques will help us to understand what it takes to be a successful invader by identifying drivers of invasive potential and species response to changing environments.”

The researchers hope that the genomics project will lead to tools that Biosecurity officers can use at the border and in the field, if a pest invasion threat is likely or a new pest is discovered, to better understand the biology of the invader and inform the management strategy needed. “This is why end-user involvement at the beginning of this work is so important.” says Manpreet.

A typical border interception scenario would involve sequencing the genome of the intercepted species (whether known or unknown) and running the resulting sequence through the Invasomics pipeline. The pipeline would compare the genomic traits of the individual with the predefined matrix of genomic characteristics associated with invasiveness to generate an invasiveness score (a metric of how serious a biosecurity threat it might be). The invasiveness score would then be used in parallel with other methods in the MPI toolkit, such as climate analysis and pathway analysis, to determine the appropriate response.

A special Society for Molecular Biology and Evolution meeting on Invasomics is planned later in 2022, to continue developing this relatively new field of research. And the world’s first Invasomics Hub highlights the outstanding early leadership of Genomics Aotearoa researchers in this emerging field.

PROJECT LEADS:

Prof Thomas Buckley (Manaaki Whenua – Landcare Research)
Dr Shannon Clarke (AgResearch)
Dr David Chagné (Plant & Food Research)
Dr Anna Santure (University of Auckland)

High Quality Genomes and Population Genomics for Primary Production

Primary production end-users can use high quality genomes and population genomics tools to better predict biological traits such as the reproduction, appearance, behaviour, or disease susceptibility of a variant, and how these traits differ among populations. Understanding this variation is used to improve economic yields and increase resilience to environmental change.

End-users have benefitted from the project output through adoption of genomic tools to solve problems. For industries that already have genomics embedded in their research this means faster and more efficient analysis and deeper knowledge gained from genomic data.

For sectors that have not started to use genomics, it means removing the technical barriers to adoption that previously existed. For example, the fisheries industry in Aotearoa now has deeper knowledge of the genetic structure of hoki thanks to the adoption of new genomics technologies. Practically, the development of a high-quality genome and analysis of populations of hoki from around NZ has been achieved rapidly thanks to Genomics Aotearoa's new solutions and newly built capability.

Long read high fidelity Pacbio technology (described earlier in the Environment theme for this project) has also had an impact on primary production and enables the team to unravel complex genomes, such as the polyploid genomes of blueberry and kiwifruit.

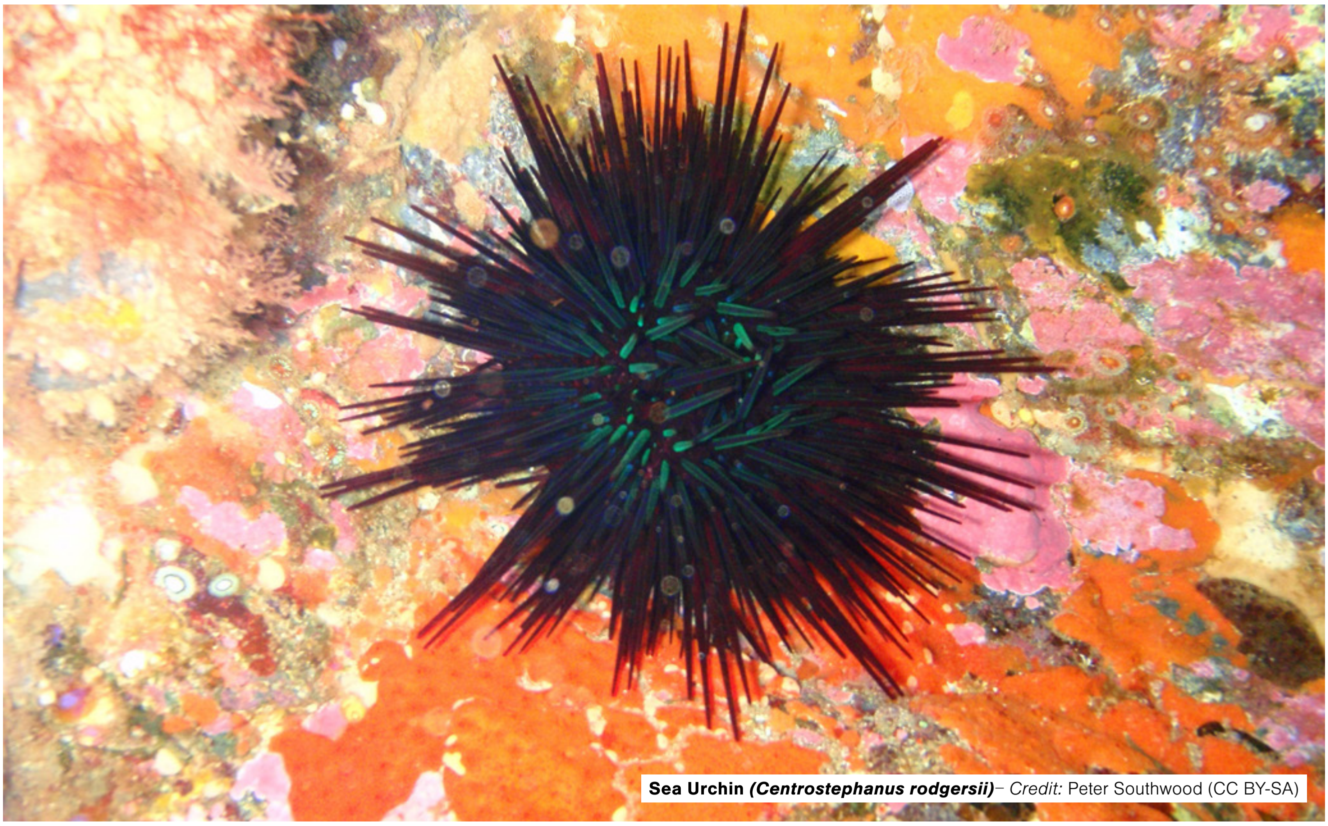
A key technical highlight for the year is the automated pipeline for high quality genome assembly developed in collaboration between Susan Thomson, Chen Wu and David Chagné (Plant & Food Research) and INRAE UMR BIOGECO, France and co-funding from Plant & Food Research's kiwifruit breeding programme. This pipeline was tested using a tetraploid kiwifruit female, male and two offspring quartet sequenced using PacBio. The automated pipeline ran for 35 hours – a very short time considering the complexity of the analysis. The resulting genome assembly was of high quality and all four sets of chromosomes were resolved, nearly all telomeres detected, and the completeness of the assembly was close to 99%. This achievement is a technical breakthrough, enabling the application of the same pipeline to other complex polyploid genomes. The resulting high-quality genome is being used to determine the genetic control of key traits for kiwifruit breeding.

Developing new aquaculture initiatives in collaboration with Te Whānau-a-Apanui

Te Whānau-a-Apanui hapū in the Eastern Bay of Plenty and East Cape is remote from many scientific and research institutions, but the hapū has a commitment to regional development in marine aquaculture, as well as informed stewardship of their coastal marine ecosystems. Two new projects are being built with their involvement from the ground up, with the additional benefit of the co-leadership of a researcher of Te Whānau-a-Apanui descent.

Kelp (*Ecklonia radiata*) and the long-spined urchin (*Centrostephanus rodgersii*) are not currently subject to aquaculture, but both species have potentially high market values, domestically and internationally.

Brown seaweed, which includes *Ecklonia radiata*, makes up around one-third of the international seaweed market. This market is projected to grow from USD 15.01 billion to USD 24.92 billion by 2028. Some Asian and European markets are showing increasing demand for highly prized urchin species. Aotearoa's endemic kina does not attract the high prices of other urchins. The long-spined urchin is also found in Australian waters, where a small, but growing aquaculture industry exists, indicating some potential for the species in Aotearoa also.



Sea Urchin *(Centrostephanus rodgersii)* – *Credit:* Peter Southwood (CC BY-SA)

J. David Aguirre (Massey University) is of Te Whānau-a-Apanui descent and is committed to supporting this aquaculture development. The Project team worked with the Te Whānau-a-Apanui Trust Board to decide on sampling locations and talk about the research practice and data management plan. It was decided that in August 2022, a hui would discuss the data generated in the project (and similar projects), and its appropriate stewardship. Several national leaders in genomic data sovereignty and representatives from surrounding hapū and iwi have been invited.

Assisted by a Rutherford Discovery Fellowship to J. David Aguirre, the generation of a genome by the Massey University team (J. David Aguirre, Vanessa Arranz, Libby Liggins) in collaboration with Plant and Food (David Chagné, J. Ignacio Carvajal, and Elena Hilario) is helping develop culturing and aquaculture of kelp. The outcomes will aid ecological restoration and farming efforts important to Te Whānau-a-Apanui, other coastal communities, and the marine industry.

The kelp sub-project is closely aligned to the long-spined urchin genome sub-project co-led by Libby Liggins (Massey University). Libby is also funded by a Rutherford Discovery Fellowship. The generation of a genome in collaboration with Annabel Whibley and Auckland Genomics, University of Auckland supports population genomic investigations into the historical and contemporary range dynamics of the species in Aotearoa.

Collaboration with industry and international experts enhances research

The Genomics Aotearoa Plant & Food Research team members are part of a collaboration with CNRS, France focusing on the role of transposable elements in the genome of fruit tree crops during domestication, and capitalising on the *Gillenia* genome assembly work. The team are also involved in an international collaboration with USDA-SCRI (the VacCAP project) to develop a pangenome for *Vaccinium* (blueberry), building tetraploid genome assemblies for blueberry, and developing genomics tools useful for blueberry breeding, building on the earlier Bilberry assembly work reported last year.

Researchers Shannon Clarke and Rudiger Brauning co-lead the Ovine PanGenome project, a USDA funded international collaboration, which is aligned to the Genomics Aotearoa HQG+PG and Genome Graph projects. Another pangenome collaboration aligned to HQG+PG is the genome assembly of economically important ryegrass with the International Lolium-Festuca Pangenome Consortium.

Like the Environment theme sub-projects, the number of publications resulting from the HQG+PG project in the primary production sector reflects both the effort and high quality of the research programme.[2]

---

[2] Fang Z-Z et al. 2022. The genome of low-chill Chinese plum “Sanyueli” (*Prunus salicina* Lindl.) provides insights into the regulation of the chilling requirement of flower buds. Molecular Ecology Resources 22(5): 1919-1938. doi.org/10.1111/1755-0998.13585

Ireland HS et al. 2021. The *Gillenia trifoliata* genome reveals dynamics correlated with growth and reproduction in Rosaceae. Horticulture Research 8(1): 233. doi.org/10.1038/s41438-021-00662-4

Koot E et al. 2021. Genome-wide analysis reveals the genetic stock structure of hoki (*Macruronus novaezelandiae*). Evolutionary Applications 14(12): 2848-2863. doi.org/10.1111/eva.13317

Li Z et al. 2022. Chromosome-scale reference genome provides insights into the genetic origin and grafting-mediated stress tolerance of *Malus prunifolia*. Plant Biotechnology Journal 20(6): 1015-1017. doi.org/10.1111/pbi.13817

Tahir J et al. 2022. First Chromosome-Scale Assembly and Deep Floral-Bud Transcriptome of a Male Kiwifruit. Frontiers in Genetics 13. doi.org/10.3389/fgene.2022.852161

Tian Y et al. 2022. Transposon insertions regulate genome-wide allele-specific expression and underpin flower colour variations in apple (*Malus* spp.). Plant Biotechnology Journal 20(7): 1285-1297. doi.org/10.1111/pbi.13806

Wu C et al. 2022. A chromosome-scale assembly of the bilberry genome identifies a complex locus controlling berry anthocyanin composition. Molecular Ecology Resources 22(1): 345-360. doi.org/10.1111/1755-0998.13467

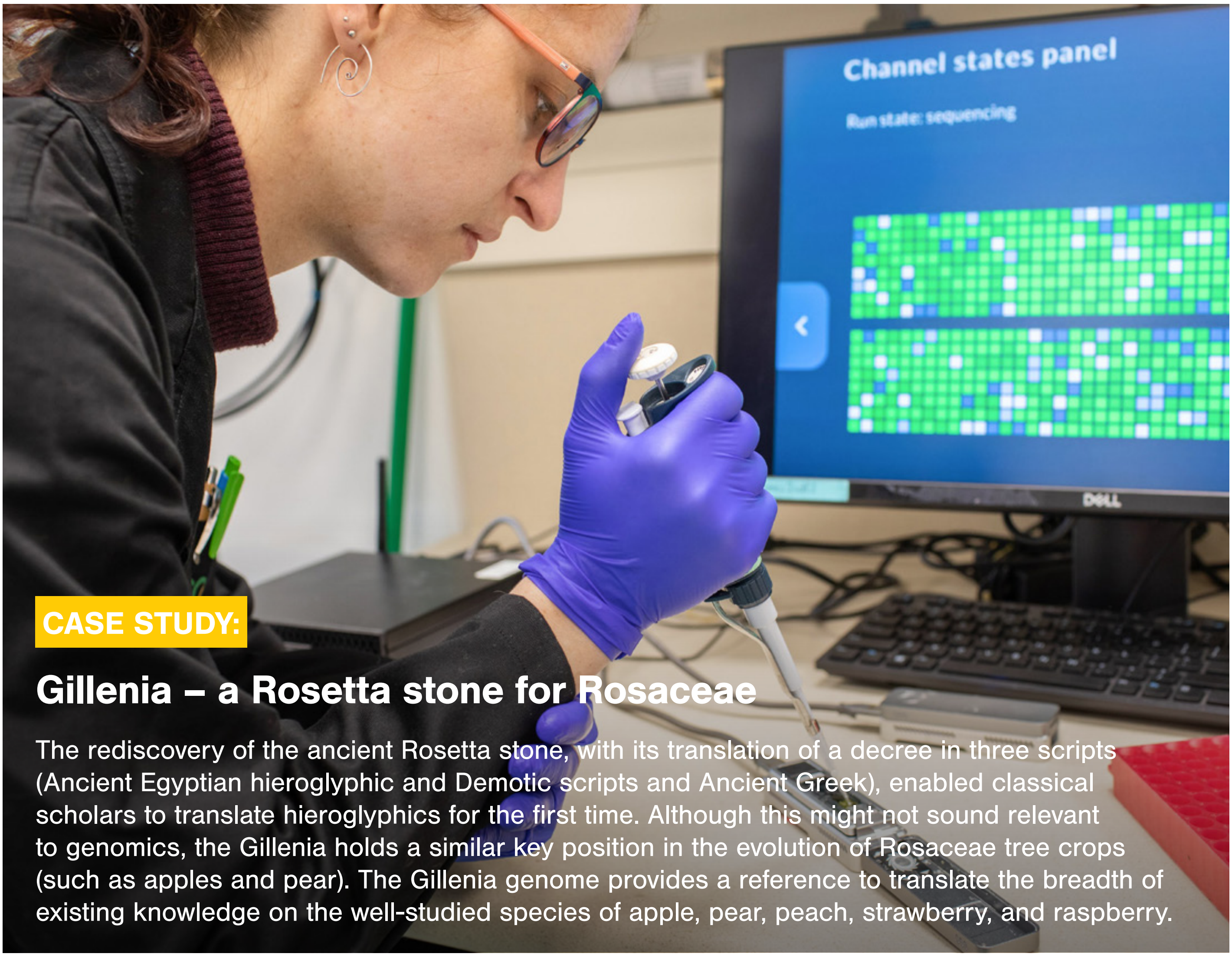
Hess A et al. 2022. Expanding the genomic tool kit: What does Oxford Nanopore Sequencing have to offer? accepted into the Proceedings of the 12 World Congress on Genetics Applied to Livestock Production. (peer-reviewed conference proceeding)

Completed Projects Continue to Provide Industry Benefits

Methods and know-how generated in the Better Breeding Values project continue to contribute to research and help to improve genetic gain for the sheep and cattle industries.

The infrastructure used as part of the genomic evaluations for the Aotearoa sheep industry leveraged outcomes from the Better Breeding Values project to provider more accurate imputation, ultimately contributing to improved genetic gain.

The sheep and cattle industries continue to generate large volumes of genomic data. The analysis pipelines and know-how from the Better Breeding Values project continue to support the analysis of these data to allow us to better understand the relationship between genotype and phenotype. This knowledge will be used to improve genetic gain, hence profitability, for sheep and cattle operations.



CASE STUDY:

Gillenia – a Rosetta stone for Rosaceae

The rediscovery of the ancient Rosetta stone, with its translation of a decree in three scripts (Ancient Egyptian hieroglyphic and Demotic scripts and Ancient Greek), enabled classical scholars to translate hieroglyphics for the first time. Although this might not sound relevant to genomics, the Gillenia holds a similar key position in the evolution of Rosaceae tree crops (such as apples and pear). The Gillenia genome provides a reference to translate the breadth of existing knowledge on the well-studied species of apple, pear, peach, strawberry, and raspberry.

Last year we reported on the significance of sequencing the *Gillenia* genome as an evolutionary model for the Rosaceae, and work is progressing to unravel its secrets.

The 2021 assembly confirmed that *Gillenia* has an unduplicated genome. Genome duplication, where an organism inherits two copies of the genome of its parents, instead of only one copy, is relatively common in plant genome evolution. Duplication happened at some stage during the evolution of the family that includes *Gillenia*, apple and pear.

“Genome duplication is an important source of novelty in evolution, providing new genetic material for selection, the result of which is specialised or new gene functions. Without such gene duplication the opportunity for species to cope with changing environments would be severely limited,” says David Chagné from Plant & Food Research.

Plant & Food Research researchers led by David Chagné have produced a new genome assembly using Pacbio sequencing technology. They collaborated with Dr Amandine Cornille from CNRS, France who studies the effect of transposable elements (DNA sequences that can change their position within a genome) on fruit domestication.

The team have now sequenced genomes of domesticated and wild representatives of apricot, apple, pear, and almond to determine if transposable elements have a role in characteristics selected during domestication.

Gillenia is being established as a model to further understand the critical biological processes that coordinate cellular activity (known as transcriptional regulation) of fruit flesh development in apple.

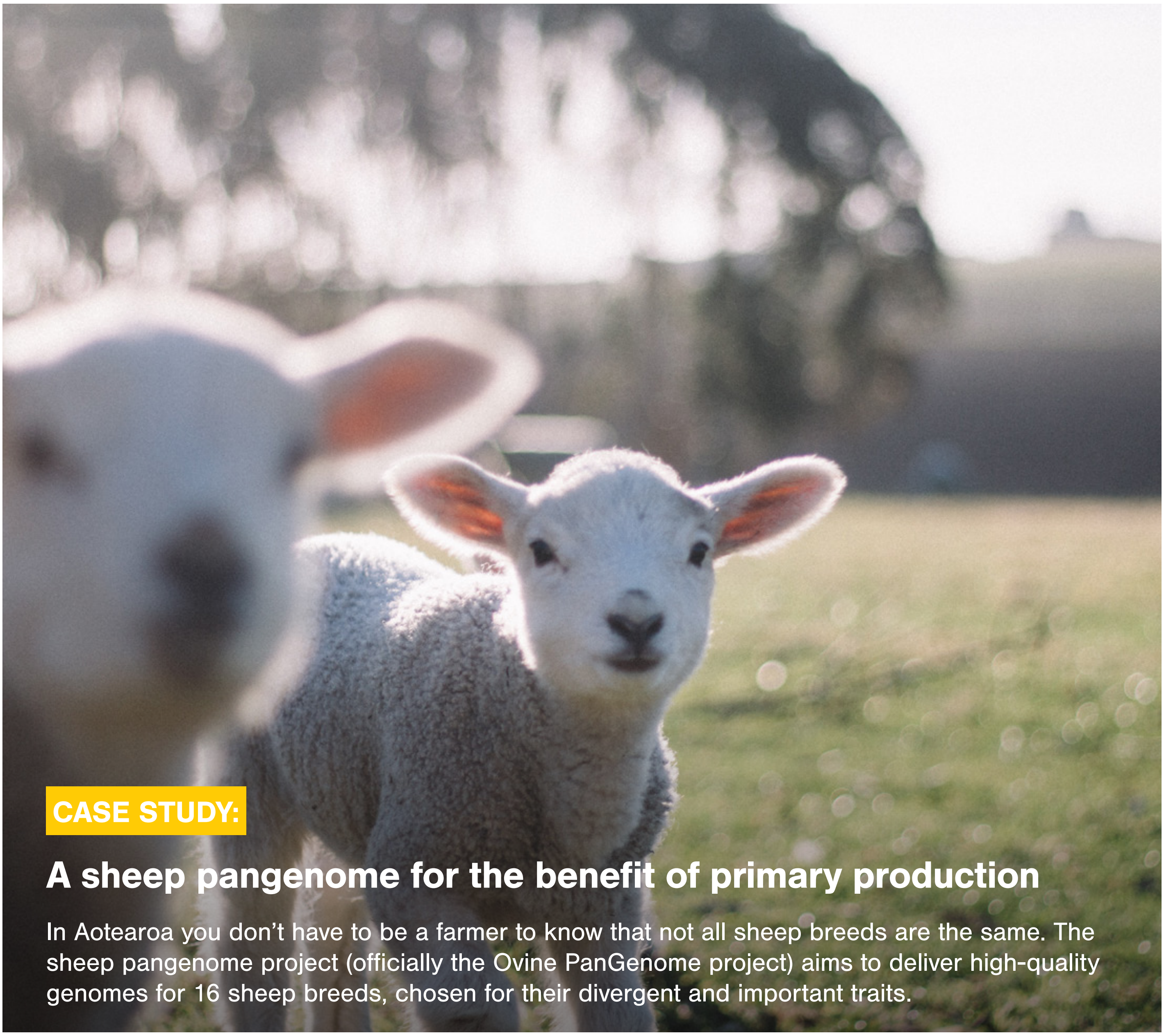
“We have fresh insight into differences and what is shared between fleshy and dry fruit development. Understanding the key drivers of fleshy, tasty fruit has enormous potential for further developments for our tree crop production in Aotearoa,” David says.

As an unduplicated genome, *Gillenia* holds a key position in the evolution of Rosaceae tree crops, providing the means to compare genomes across the entire Rosaceae plant family. *Gillenia*’s growth, flowering and fruiting characteristics can be mapped among species in the family.[3]

The knowledge gained through this work has enormous potential for further developments for our tree crop production, with global benefits.

---

[3] Ireland HS, Wu C, Deng CH et al. 2021. The *Gillenia trifoliata* genome reveals dynamics correlated with growth and reproduction in Rosaceae. Horticultural Research 8: 233. doi.org/10.1038/s41438-021-00662-4



CASE STUDY:

A sheep pangenome for the benefit of primary production

In Aotearoa you don't have to be a farmer to know that not all sheep breeds are the same. The sheep pangenome project (officially the Ovine PanGenome project) aims to deliver high-quality genomes for 16 sheep breeds, chosen for their divergent and important traits.

The project aims to deliver a more detailed understanding of the genes that influence productivity in sheep, in turn helping the sheep industry in Aotearoa stay ahead of future needs for meat and wool.

The sheep pangenome is an international collaboration headed by Dr Brenda Murdoch from Idaho University and funded by the USDA-NIFA. Aotearoa is represented by Shannon Clarke and Rudiger Brauning from AgResearch.

The project has direct local industry input and alignment. The mating, raising and subsequent tissue collection (parents and offspring) were provided as in-kind contributions by local industry partners.

So far, the project has identified genomic regions associated with a range of phenotypically distinct and important traits in different breeds. The assembly of multiple genomes overlaid with information from phenotype datasets (such as transcription and methylation profiling), enables the researchers to compare genomic regions associated with traits of importance.

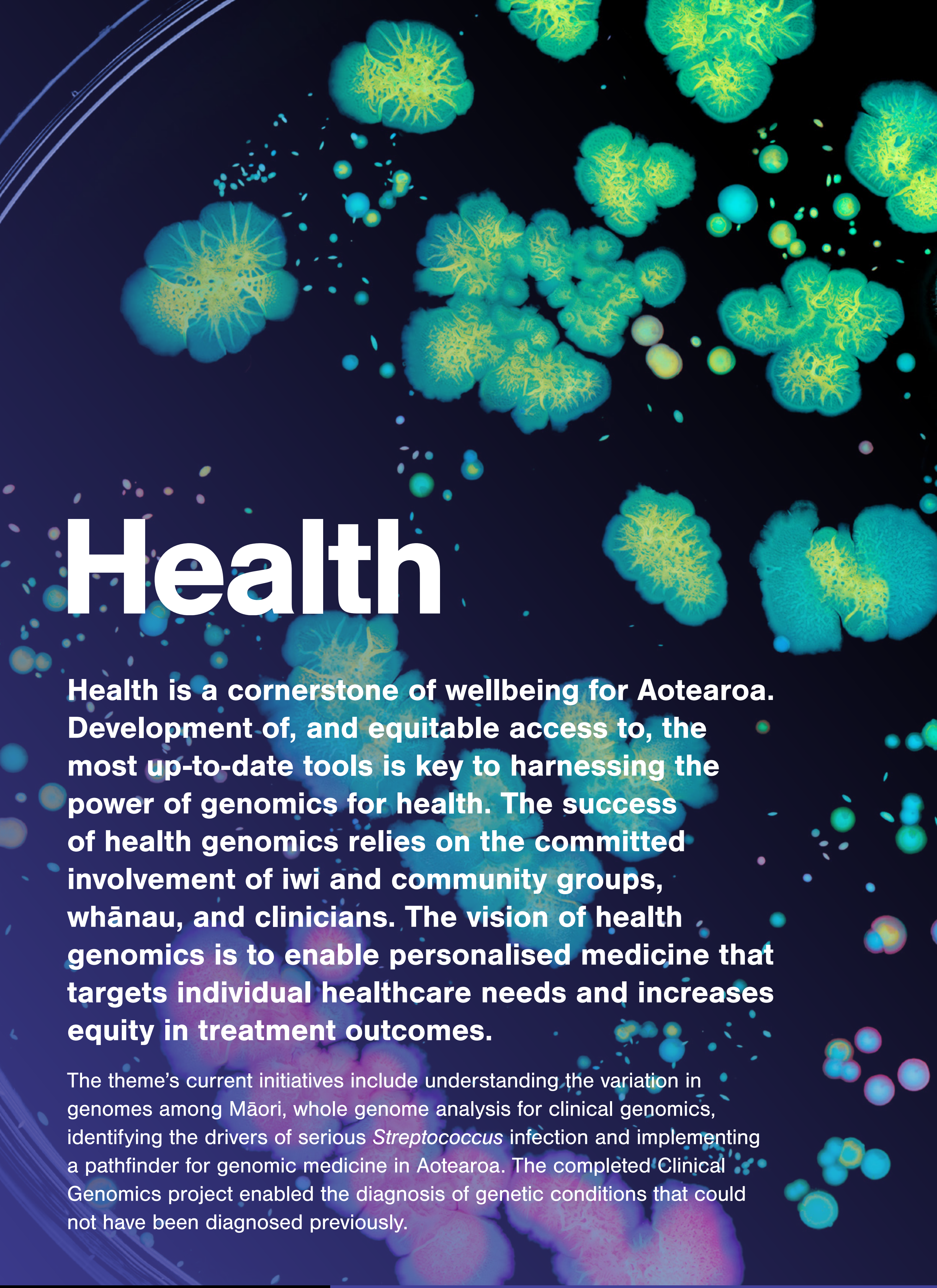
Shannon says, "this capability will allow us to fill in the gaps in genomic knowledge, to analyse genetic variants to understand traits and take full advantage of the diversity within local sheep breeds."

Pangenome assemblies, alongside an increase in high-quality reference genomes with functional annotation, provide fundamental resources needed for industry to use functional genomics as a tool.

The sheep pangenome project supports more accurate prediction of traits and improved breeding strategies.

"It's about understanding the genomic architecture behind phenotype, so we can take the breed in whatever direction we want to go. This enables more accurate selection now, but also helps answer questions for the future, such as how breeds will adapt to a changing environment – how to improve breed resilience, and how to maintain production under the predictions of climate change," says Shannon.

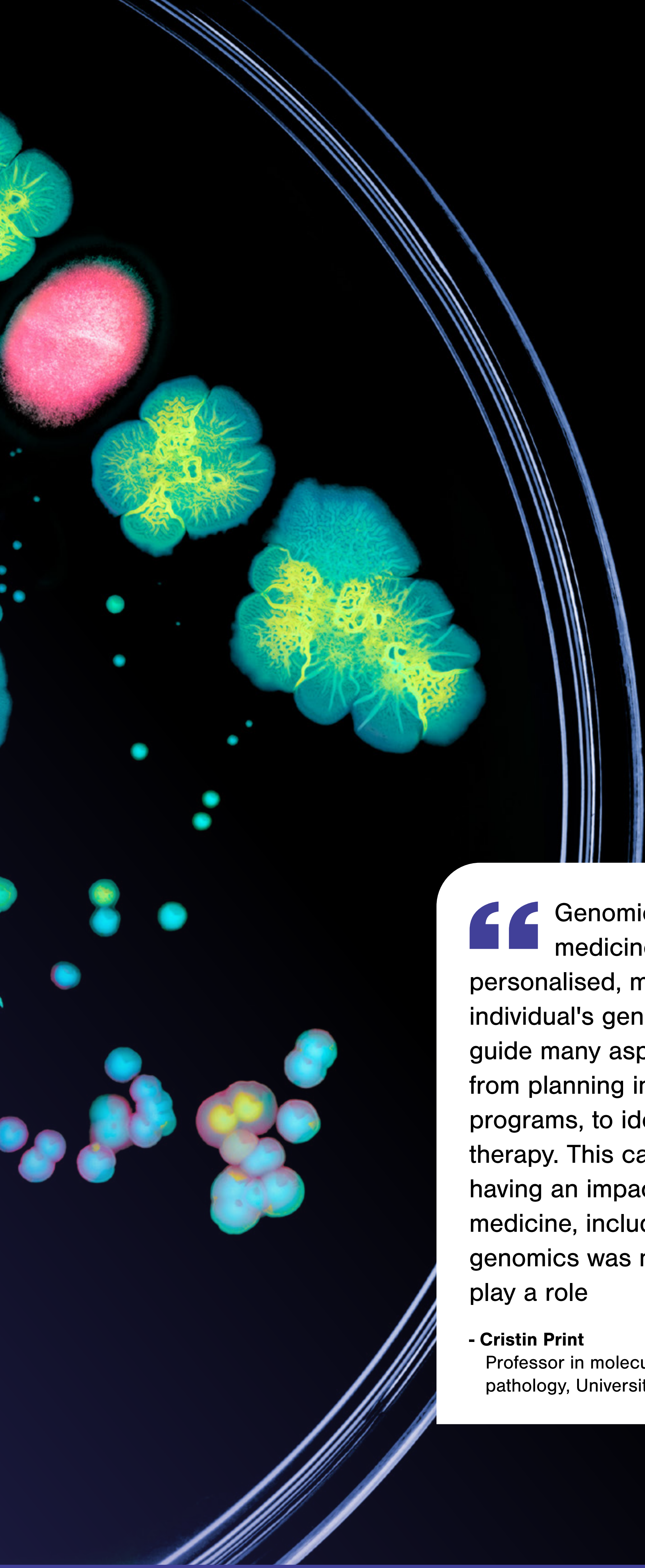
The sheep pangenome will also contribute to comparative studies among species, while complementing and leveraging the pangenomes being constructed for cattle, and other ruminants.



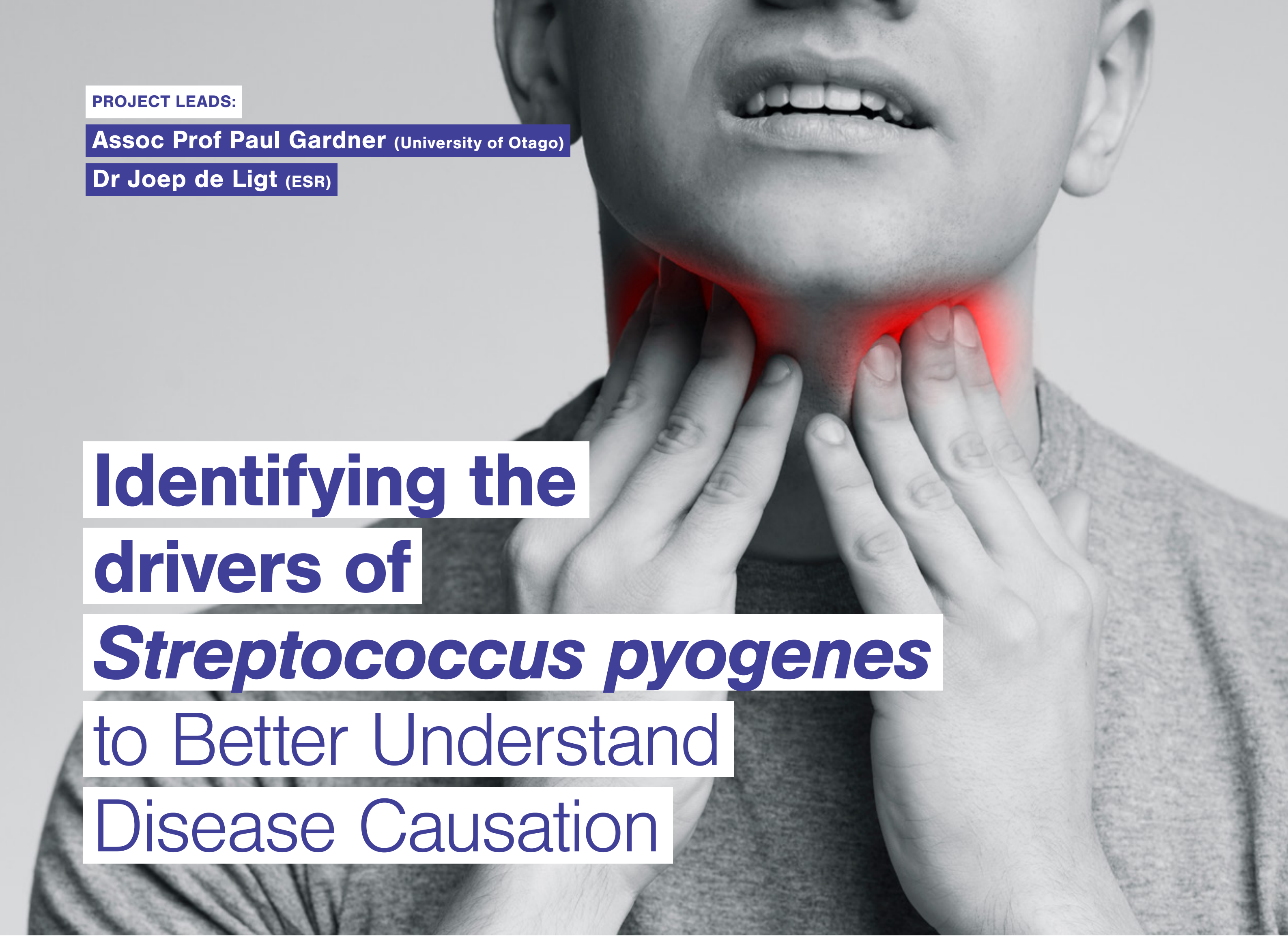
Health

Health is a cornerstone of wellbeing for Aotearoa. Development of, and equitable access to, the most up-to-date tools is key to harnessing the power of genomics for health. The success of health genomics relies on the committed involvement of iwi and community groups, whānau, and clinicians. The vision of health genomics is to enable personalised medicine that targets individual healthcare needs and increases equity in treatment outcomes.

The theme's current initiatives include understanding the variation in genomes among Māori, whole genome analysis for clinical genomics, identifying the drivers of serious *Streptococcus* infection and implementing a pathfinder for genomic medicine in Aotearoa. The completed Clinical Genomics project enabled the diagnosis of genetic conditions that could not have been diagnosed previously.



> Genomics is transforming medicine. Precision, or personalised, medicine uses an individual's genomic information to help guide many aspects of their healthcare, from planning individualised screening programs, to identifying the best drug therapy. This capability is progressively having an impact across all areas of medicine, including many diseases where genomics was not previously thought to play a role
>
> **- Cristin Print**
> Professor in molecular medicine and
> pathology, University of Auckland



PROJECT LEADS:

Assoc Prof Paul Gardner (University of Otago)

Dr Joep de Ligt (ESR)

Identifying the drivers of *Streptococcus pyogenes* to Better Understand Disease Causation

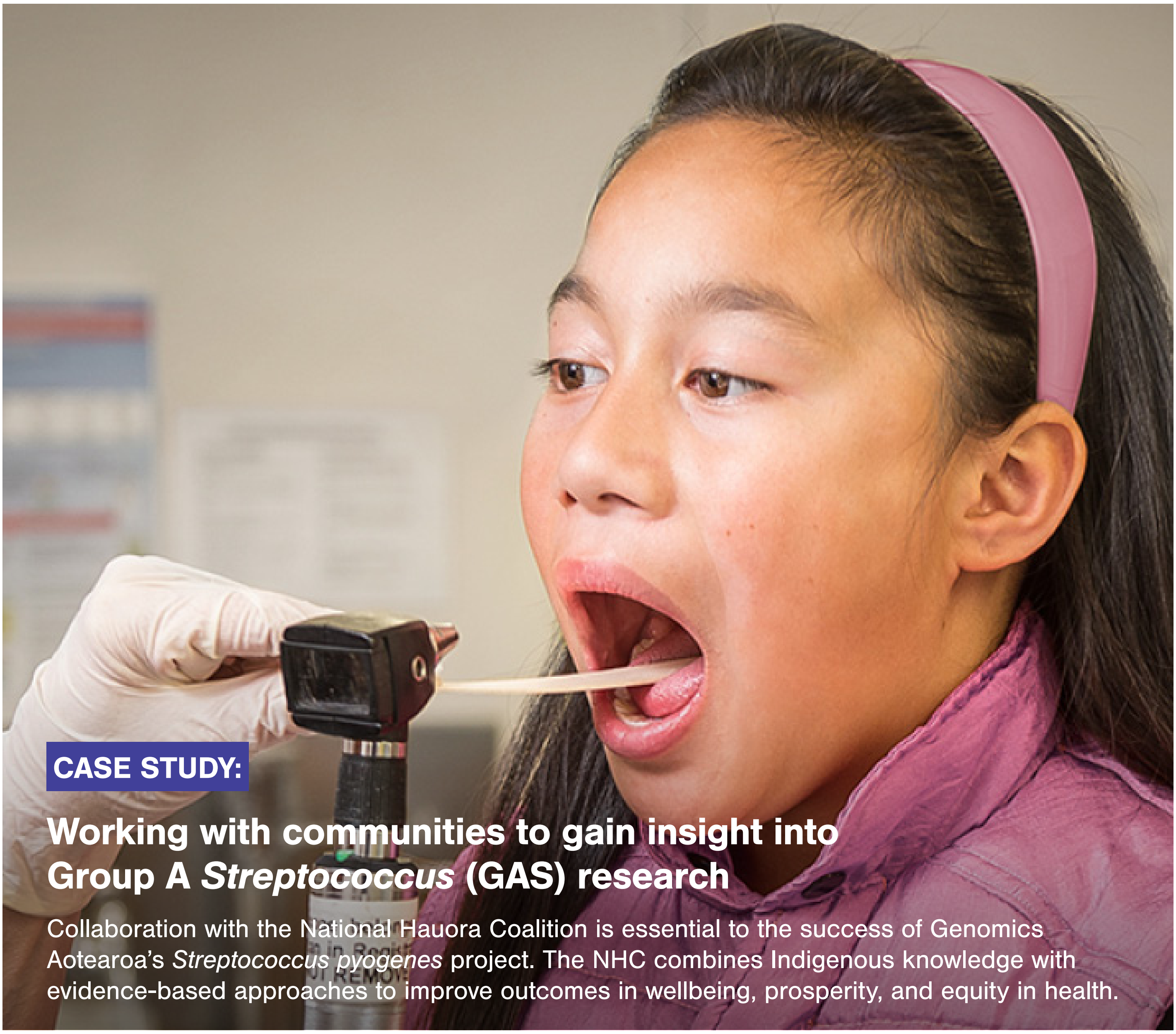
Globally, more than 600 million Group A Streptococcus (GAS; *Streptococcus pyogenes*) infections occur each year. Māori and Pasifika are disproportionately affected by GAS, including many children. Although GAS infection can result in a range of clinical outcomes, from minor sore throats and skin infections to rheumatic fever and rheumatic heart disease, we still do not fully understand the genetic drivers associated with different outcomes.

In bacterial infections, the relationships between genotype (genetic causes) and phenotype (invasiveness or resistance to antimicrobial agents) are difficult to predict from genome sequences. To better understand the causes of the range of GAS outcomes, the project aims to 1) gain insights into the genetic diversity and population structure of invasive and non-invasive GAS circulating amongst children in Aotearoa and 2) identify genetic drivers responsible for invasive phenotypes (genotype-phenotype prediction).

The project uses a machine learning algorithm to identify genetic variants that are associated with more invasive bacteria (those that are more likely to cause severe disease), which requires a large and diverse dataset to *train* the machine. Due to the large dataset already available for *Campylobacter jejuni*, this was used as a proof-of-concept of the machine learning approach.

During this first year of the project, the team sub-sampled the community GAS sample collection stored with Labtests (1,500 of approximately 14,000 samples), to identify the key strains for the project to focus on. The generation of sequencing data is currently in progress. The machine-learning approach will be applied to the resulting sequences. The team plans to share the technical skills needed when using a machine learning approach on sequence data, through a planned workshop.

The team is working with the National Hauora Coalition (NHC), a Māori-led charitable trust, on the community engagement initiative. This community engagement initiative is designed and led by Māori to gain Pasifika and Māori community feedback on the use of existing collections for further research.



CASE STUDY:

Working with communities to gain insight into Group A *Streptococcus* (GAS) research

Collaboration with the National Hauora Coalition is essential to the success of Genomics Aotearoa's *Streptococcus pyogenes* project. The NHC combines Indigenous knowledge with evidence-based approaches to improve outcomes in wellbeing, prosperity, and equity in health.

The partnership with NHC is important to ensure Māori- and Pasifika-informed culturally responsive processes and data sovereignty are embedded in the project. Ultimately, NHC will be part of an initiative for improving health outcomes for Māori and all populations in Aotearoa.

Tira Phillipson-Puna, a research assistant with the NHC, has been leading the community consultation process, seeking insights from Māori, Cook Island Māori, Tongan and Samoan communities. Avoiding a one-size-fits-all approach is important. Considering different ethnic, iwi and regional perspectives ensures genuine rather than tokenistic inclusion in research.

"These communities provided the original samples, so we wanted to engage fully with them to get their feedback on using the samples for this project. It's important to know how they feel about the researchers using existing collections for research, what that means for data governance, and how to share information and results from the project back to those involved," Tira says.

Tira has made significant progress in the engagement process, using recruitment advertisements to attract both healthcare worker and patient participants through the NHC networks. The advertisements were followed by two pilot workshops and two focus groups, with additional workshops planned for patient participants.

The workshops focused on awareness, information sharing, and involving participants in the process. The primary question posed by the consultation was: "How should a laboratory care for your *Streptococcus* sample throughout its research journey?"

Information resources described what *Streptococcus* is, the disease it can cause, the process for taking samples, how the samples would be used, and the journey of the samples through the laboratory process.

The initial workshops helped discover the level of information needed by the participants. As a result, future workshops will simplify some of the more specialised terminology and tailor the language used to suit both healthcare workers and patients.

Ultimately, the outcomes of this engagement will form the basis of guidelines for researchers who regularly collect samples or already have collections of samples from Māori or Pasifika communities.

PROJECT LEADS:

Prof Stephen Robertson and Assoc Prof Phil Wilcox (University of Otago)

The Aotearoa New Zealand Genomic Variome:

Capturing Genomic Diversity for Disease Dagnosis

Variomes describe the unique genetic signatures of populations. Without a comprehensive variome for Aotearoa, shaped as it is by our unique human history, genomic tools based on overseas non-Māori populations have limited use in genomic medicine for the increasing number of people with Māori ancestry.

The Variome project categorises Māori genomic variation, to improve disease diagnosis and to support healthcare research driven by Māori, for Māori. Genomic data sourced from elsewhere are not always informative for Māori. Sometimes variants associated with disease susceptibility in non-Māori populations have no such relationship in Māori, and sometimes associations are completely unknown. The Variome project seeks to address this health inequity.

Critical to this project is that it is Māori-led, with co-development that addresses ethical use and recognises data sovereignty, developing processes to ensure appropriate consenting. A leadership rōpū ensures that research priorities are co-developed for the benefit of communities and that benefits arising from the data and kaitiakitanga are consistent with Te Ao Māori. The establishment of tikanga and control and governance within Māori hands ensures data sovereignty.

The Variome project continues to assemble the catalogue of variation in the genomes of people with Māori ancestry, to ensure that genomic medicine for Māori can be accurately applied. Understanding the variome of Māori is critical for effective, accurate, and equitable genomic diagnosis in Aotearoa.

The project made significant progress this year. Additional communities have engaged with the research and by the end of the year 210 genomes had been sequenced, aligned and variants identified. A further 250 samples have been provided by participants for sequencing. Partners include Ngāti Porou Hauora (Tairawhiti), Tuiora Health (Taranaki New Plymouth), the Centre for Health (Tauranga Moana), Poutini Waiora (Te Tai Poutini), and Hāpai te Hauora (Ōpanuku) and Hāpai Te Hauora (Tāmaki Makaurau Auckland).

A key future benefit of this project is that it dovetails into both the Rakeiora and Whole Genome Analysis (clinical diagnostics) projects. Clinicians will be able to use the variome resource base to improve the accuracy and precision of genomic diagnostics.

CASE STUDY:

Māori-led recruiting of study participants is essential

When engaging with communities, recruiting Māori participants is a fundamental step.

Cinnamon Lindsay-Latimer, a Māori researcher for Hāpai Te Hauora, which holds the regional Māori public health contract for Tāmaki Makaurau Auckland, has developed a successful strategy that has boosted the number of participants in the project, and is also helping Māori understand the importance of the mahi.

Cinnamon has attended events around Tāmaki Makaurau, including those celebrating Matariki, to talk to people and ask for participants, with the aim of building trust, acknowledging the Te Ao Māori view on DNA as taonga, demystifying genomics and increasing understanding of the potential benefits for Māori health.

“We’ve had considerable buy-in, particularly from young people; I think because we’re layering this with tikanga. Volunteers trust the organisations backing the project, they are reassured the University of Otago has processes that keeps their information safe and are impressed that a karakia accompanies their DNA sample. It makes sense to them. They see the real-life examples in our handout where it has taken years to trace and treat genetic problems in a family, when a Variome will mean understanding the heritability of a disorder could take only a matter of weeks,” says Cinnamon.

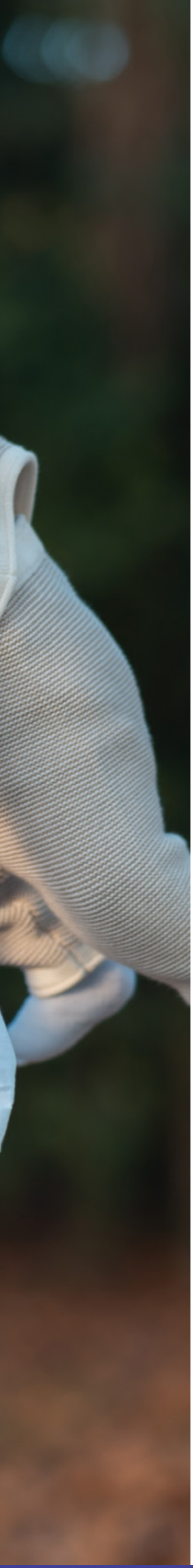
Cinnamon adds, “the lovely thing is that the participants go home to their whānau excited about being involved in something that will make a difference, and the whole whānau want to sign up. Unfortunately, they can’t because we want to get as diverse a representation of the population as possible. But it shows its being talked about – it is interpreting the science, and it’s moving us beyond historical mamae in this space.”

“They see good in it, and they see that their huge gift of DNA is respected. I’ve been blown away by the positive responses. It’s a project I’ve loved being part of.”

PROJECT LEADS:

Prof Stephen Robertson and Assoc Prof Louise Bicknell (University of Otago)

Extending Whole Genome Analysis into Healthcare



Sequencing and analysing a person's entire genome – whole genome analysis (WGA) – is the most comprehensive method of arriving at a genetic diagnosis. In Aotearoa the capability to implement WGA in healthcare has yet to be realised.

New this year, the project will work with whānau affected by an undiagnosed (potentially) genetic condition. Whānau members will have their genomes sequenced and the data analysed by specialist teams within hospitals in Aotearoa.

Building on the Genomics Aotearoa Clinical Genomics project outcomes, which established links across research and clinical teams throughout Aotearoa, this project extends diagnostics capability to a broader range of potential genetic conditions.

The project will enable much- needed exposure to diagnostics using whole genome datasets among clinicians in the Auckland, Canterbury and Wellington regions.

Whereas the Clinical Genomics project studied parent-offspring trios (in which both parents and at least one child are affected by a particular genetic condition), this project studies singletons (without data from the parents). The project develops new methods for estimating data in these situations.

The project is committed to increasing the capability of clinicians in genomics diagnosis, starting at the undergraduate level. The team has developed a Genomics Wānanga – using a Te Ao Māori approach to teach genomics analysis to Māori medical students. The first wānanga was scheduled for July 2022.

The project will enable much-needed exposure to diagnostics using whole genome datasets among clinicians in the Auckland, Canterbury and Wellington regions.

New bioinformatic tools have been tested. The project has developed a new pipeline enabling alignment of genomes to GRCh38 human reference genome and incorporated Deep Variant as an alternative to GATK-based variant calling.

So far, the team have diagnosed a genetic condition for one of the 24 people from 12 whānau involved in the study. The project also plans to recruit more whānau. Recruitment by clinicians is purposefully directed towards encouraging Māori to participate. As a result, more than half of people recruited to date self-identify as Māori. Owing to this strategy, several whānau will have a diagnosis not otherwise achievable using the public health system in Aotearoa.

PROJECT LEADS:

Prof Cris Print (University of Auckland)

Prof Stephen Robertson and Assoc Prof Phil Wilcox (University of Otago)

Rakeiora: A Pathfinder for Genomic Medicine in Aotearoa New Zealand

Rakeiora aims to transform Aotearoa's translational genomic health research infrastructure. Historically we have had high quality but siloed projects with little potential for collaboration and inconsistent inclusion of Māori Treaty rights and Māori & Pacific health needs. We need an effective, safe, adaptable, and equitable Aotearoa-wide genomic research infrastructure. That is Rakeiora's vision. The project will link genomic data and health metadata, to enable research that is connected and scalable, and which includes co-innovation with Māori and is grounded in Te Tiriti o Waitangi and tikanga frameworks.

The Rakeiora programme approach includes deliberate co-design, co-governance, data sovereignty, consenting and co-innovation with Māori. The project incorporates mātauranga Māori and has begun implementing the key principles of Māori governance over Māori genomic resources and Māori genomic (and health) data. This research framework is a unique, deliberate attempt to generate a safe, mana-enhancing and health-enabling infrastructure for Māori.

The gift of mātauranga by Māori participants is likely to generate a far better infrastructure for all than would be otherwise possible. It will prepare for the incorporation of mātauranga whakapapa to improve future genomic prediction accuracy.

Rakeiora uses action learning to explore options for the acquisition, use, and storage of genomic datasets linked to clinical information for healthcare research in two sub-projects - Primary Care (community healthcare led by a nurse or GP) and Tertiary Care (for advanced cancer treatment in a specialist facility).

The team reported last year that this year would be a busy year. That prediction has come true, and progress has been complicated by the COVID-19 pandemic and health service exhaustion. Despite the difficult context and delays, considerable progress has been made. Highlights include:

- Hui with Ngāti Porou mana whenua around the design of, and principles supporting, precision medicine research and how it can reflect tikanga Māori attributes.
- Primary Care recruitment is complete and rapid Tertiary Care recruitment is underway.
- Completion of the Ira Tātai Whakaheke standards for genomic research in cancer with Māori. This transformative document is being evaluated by several relevant groups before publication.
- NeSI initial assembly of the scalable computational framework, aspects of which have been evaluated on several occasions by Māori and non-Māori team members. This framework continues to evolve following an action learning path that records its progress and insights.

Due to the pandemic delays, the project received a time-only extension until March 31, 2023, to complete all project components.

Rakeiora has enabled considerable development in both computational infrastructure capability and in Te Ao Māori engagement for partner organisations NeSI and Cancer trials NZ. A Māori scholar working in the project, Kimiora Henare, has also contributed to SING Aotearoa.

The project has engaged in several end-user knowledge transfer events, including a formal hui to Ngāti Porou Hauora communities, a workshop with MBIE and Ministry of Health representatives, a worksho p at the eResearch NZ 2022 online conference, and a hybrid on-line and in-person workshop with clinicians and scientists at the Auckland Centre for Cancer Research.

CASE STUDY:

Clinical genomics analysis corrects a decades-long misdiagnosis

The Clinical Genomics project was completed during the previous reporting period, but its outputs continue to be used to diagnose uncommon conditions that could be diagnosed by other means

Getting an accurate diagnosis is a major hurdle for many people with rare conditions. When a person has a rare genetic condition that cannot be precisely diagnosed, the impacts on them and their whānau can be profound. A recent survey found that it took more than a year for over 50% of people with a rare condition to receive an accurate diagnosis, and for 20% of people it took more than 10 years.[1]

An extreme case is that of a man, now in his 40s, who was diagnosed with Prader Willi Syndrome as a baby. Prader Will Syndrome (ipwso.org/pws-information/what-is-pws/) is a rare genetic multi-system condition with wide a range of complex features. Despite having some features not consistent with the syndrome, the man was brought up following the treatment guidance of the time. That time was difficult for the man's mother.

"It would have made life easier if this test had been available when he was a baby – but despite the challenges, my son has developed into a healthy social man who is living a good life."

The Clinical Genomics research team found genetic variants within the man's genome that revealed a misdiagnosis; he really had CHD4-related syndrome. This is another complex genetic syndrome, only discovered in 2019, that is associated with developmental and speech delay.

While there is no treatment for the condition, the man and his family now have an accurate diagnosis that explains all symptoms and a better understanding of the condition. His mother was relieved.

"I needed to know in my lifetime what was wrong and why, and now I know. For me, it's removed doubt – I'm not left wondering. There's a sense of relief, and I'm pleased for my son and for his brother that they know. They will know what it is if it is ever passed on."

Clinical genomics offers the potential for an accurate diagnosis to assist with management of rare conditions. Aotearoa has now started to build this capability among clinicians, and to embed the tools and pipelines essential for clinical genomics diagnostics.

An important aspect of all Genomics Aotearoa projects is their potential to contribute to positive outcomes after project completion. Living with a rare condition is never going to be easy, but accurate, early diagnosis could influence the trajectory of the condition's outcomes.

---

[1] 2021 New Zealand Voice of Rare Disorders Survey. Medicines New Zealand. raredisorders.org.nz/assets/VOICE-OF-RARE-DISORDERS-White-Paper-February-2021-FINAL.pdf

Technical Infrastructure

to Build Bioinformatics Capability

High quality computing infrastructure and support enables the growth of genomics research in Aotearoa, for Aotearoa. World-class researcher bioinformatics skills, secure state-of-the-art computing facilities and data curation, and mentoring future bioinformatics researchers are all essential needs.

All of Genomics Aotearoa's projects are supported by the on-going development of the national bioinformatics infrastructure to grow the skills and capability of our genomics researchers. Bioinformatics infrastructure links projects together through common tools and skill sets, helps to define a structure for capability building (e.g., training) and researcher community development, and enables coordination and resourcing at a national level.

Genomics Aotearoa has established a strong foundation for on-going support and capability building in bioinformatics. These foundations include a strong leadership team of experts from several institutions, working groups that review and progress roadmaps for bioinformatics, computing and data management, and a full-time training coordinator.

> A milestone for us this year has been that we have trained more than 1,000 people – that's an awesome achievement for the team, and it's fantastic that some of our learners have gone on to become instructors themselves. This is how we build an infrastructure of people
>
> **– Mik Black**
> Professor, University of Otago

NeSI continues to deliver computing power to support projects

New Zealand eScience Infrastructure (NeSI), with its specialised infrastructure, support-staff, and computational science expertise, provides critical support for bioinformatics research in Aotearoa. The NeSI strategy to support research aligns with Genomics Aotearoa's commitment to increase the capability of our genomics researchers. This strategic alignment makes us ideal partners.

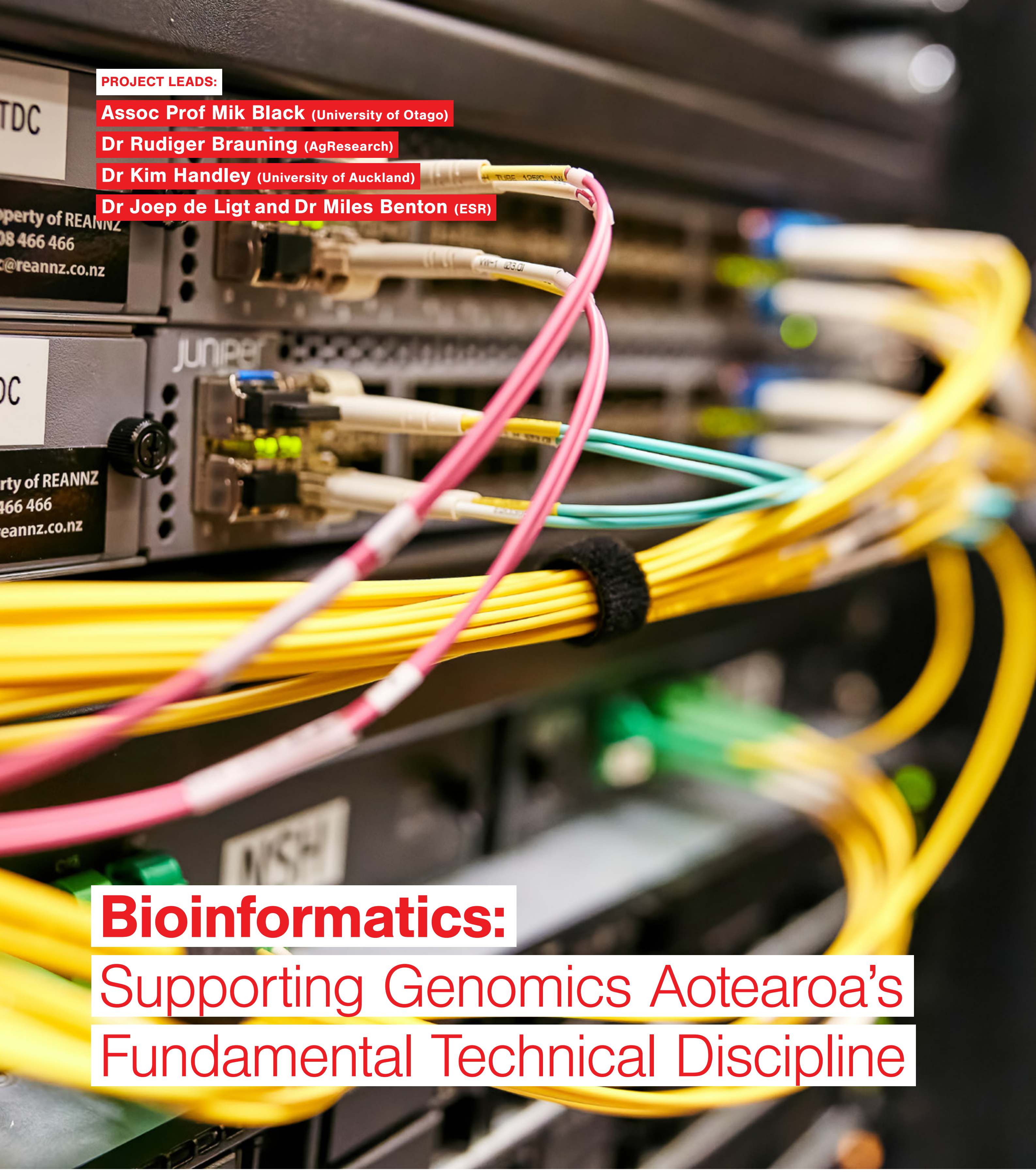
NeSI compute resources meet the needs of Genomics Aotearoa projects, support the delivery of workshop-based bioinformatics training, and host the Aotearoa Genomics Data Repository (AGDR). Genomics Aotearoa and NeSI staff work together on the development of a virtual computing environment tailored to genomics work, and the partners co-deliver bioinformatics training to the research community.

In February 2022 Genomics Aotearoa co-hosted the annual eResearch NZ Conference as an online event with NeSI and REANNZ. In 2023 GA will take on the lead host role, with a hybrid conference planned in Hamilton.

Completed projects continue to enhance bioinformatics infrastructure

several completed projects developed tools and approaches that were embedded into the bioinformatics infrastructure. Among several examples is the Epigenome-Wide Association Study tool for examining DNA methylation markers across the entire human genome (EWASP). Methylation is an epigenetic mechanism that regulates gene expression.

The EWASP tool allows researchers in Aotearoa to investigate methylation associations to improve population specific, chronic disease risk predictors. The most recent upgrades enable analysis of other species such as mice, and researchers plan to add sheep and cattle in the next 12 months. In addition, the team have developed automated techniques to identify DNA methylation patterns suggestive of ethnicity effects. This is significant as some DNA methylation sites are influenced by genetic differences among ethnic groups. This new functionality is important, as the team are committed to ensuring that DNA methylation markers work equitably, as risk prediction tools, in all population groups within Aotearoa.



PROJECT LEADS:

Assoc Prof Mik Black (University of Otago)
Dr Rudiger Brauning (AgResearch)
Dr Kim Handley (University of Auckland)
Dr Joep de Ligt and Dr Miles Benton (ESR)

Bioinformatics:
Supporting Genomics Aotearoa's Fundamental Technical Discipline

Bioinformatics combines biology, computer science, mathematics, and statistics to enable the tools and strategies needed to analyse genomic data. This is the key discipline that underpins Genomics Aotearoa's work. Although researchers arrive at bioinformatics from a variety of fields, they all share the same problem: how to analyse large, complex sets of data to extract biological meaning.

Bioinformatics combines biology, computer science, mathematics, and statistics to enable the tools and strategies needed to analyse genomic data. This is the key discipline that underpins Genomics Aotearoa's work. Although researchers arrive at bioinformatics from a variety of fields, they all share the same problem: how to analyse large, complex sets of data to extract biological meaning.

The greater the capability in bioinformatics, the more we can do with the information we gather, and the bioinformatics requirements of the research community drive our other infrastructure needs: computing resources, data management and training.

The Bioinformatics 2 project builds upon the earlier Bioinformatics project by enabling national leadership to support and develop on-going capacity and capability in Aotearoa. The project also fulfils the technical requirements to support the cultural infrastructure of Genomics Aotearoa, such as the AGDR.

NGONI AND DININDU

A strong team enables bioinformatics capability support

People, not technology, are the true drivers of Genomics Aotearoa's bioinformatics capability.

As well as the leaders, key team members include Libby Liggins (Massey), who is providing advice on the AGDR. Leah Kemp, Rachel Ashby and Michael Hoggard are supporting Joep de Ligt, Rudiger Brauning and Kim Handley respectively.

After several years of outstanding support, Genomics Aotearoa's inaugural training coordinator Ngoni Faya has left Aotearoa to pursue a clinical genomics opportunity in the United States. Our new training coordinator, Tyler McInnes, took up his position in late July, based at the University of Otago. Jian Sheng Boey is our bioinformatics trainer based at the University of Auckland.

Training continues to support bioinformatics capability development in Aotearoa

The Genomics Aotearoa/NeSI training programme is extremely popular. Twenty workshops were run this year, attracting 394 attendees. A substantial proportion of the attendees were non-Genomics Aotearoa researchers, highlighting Genomics Aotearoa's inclusive benefit of the whole research community.

The training partnership between Genomics Aotearoa and NeSI has continued to be a major strength of this project, with 52 joint workshops held over the past three years across 23 different institutions, and 1,049 attendees trained. Hitting the *1,000 trainees* milestone this year is a huge tribute to our key trainers, Ngoni Faya and Dinindu (Dini) Senanayake (NeSI) who both put an enormous amount of mahi into the workshops.

Training is also delivered directly through projects. The annual Metagenomics Summer School began during the Environmental Metagenomics project. Unfortunately, COVID-19 and the Auckland lockdown prevented the third iteration of the course being held, but it will take place again in late 2022.

The HQG+PG project team organised a statistical genomics course in Palmerston North from May 2-6, 2022. The course was delivered by Bruce Weir (University of Washington, USA) and covered theoretical and practical aspects of population genomics analysis. Thirty-four researchers from universities and CRIs attended the workshop, including 23 in-person. Emily Koot, Anna Santure, and David Chagné helped coordinate the course.

In June 2022 the HQG+PG team held a workshop for the project's early-career researchers at the University of Canterbury, co-led by Jana Wold and Tammy Steeves. The workshop included investigators across three GA projects (HQG+PG, Genome Graphs, Bioinformatics Capability) as well as investigators working on non-GA affiliated (but well-aligned) research from Plant and Food Research, AgResearch and the University of Auckland. The participants investigated tools and approaches for structural variant (SV) calling in non-model organisms using data from the HQG+PG and aligned projects. One outcome was the co-development of a project to better integrate SVs (including transposable elements) in the small population paradigm. This new project will compare the behaviour of single nucleotide polymorphisms (SNPs) and SVs across multiple generations in several species with small and/or decreasing effective population sizes. This outcome demonstrates how the project facilitates leadership by early-career researchers and leads to more research.

CASE STUDY:

The Genomics Aotearoa/NeSI training partnership goes from strength to strength

The Genomics Aotearoa/NeSI programme, run by inaugural Training Coordinator Ngoni Faya working closely with Dini Senanayake at NeSI, has grown significantly, from offering general introductory bioinformatics courses and hands-on skill-building events, to courses tailored to individual organisations. The partnership now offers a full programme with topics to suit a range of researchers – from beginners to advanced.

Ngoni says, "We have seen our capability growing from a few specialists to a community of genomics people – it started with Dini and me, and now we have an army of experts behind the two of us, enabling us to reach the 1,000-trainee milestone. Our training has developed into a one-stop-shop for those starting projects in genomic data analysis. They ask questions that either we can help with, or we can refer them to people who know the answers."

All these courses are free, and fully on-line training options have been developed. Attendees range from students to bioinformatics practitioners, managers, and researchers. As well as coming from our partner organisations, we have hosted learners from the University of Canterbury, Lincoln University, University of Melbourne, Auckland University of Technology, Unitec Institute of Technology, National Institute of Water and Atmospheric Research, GNS Science, Livestock Improvement Corporation, Museum of New Zealand Te Papa Tongarewa, Malaghan Institute, Ministry for Primary Industries, Ministry of Health, and the Cawthron Institute.

Under the direction of newly appointed Training Coordinator Tyler McInnes the programme will further expand to delivering customised software environments and data management services.

At the same time Bioinformatics Trainer Jian Sheng Boey is helping to organise and share the workflows from bioinformatics researchers via Genomics Aotearoa's GitHub site, making their pipelines and processes accessible for all.

Another achievement is that the knowledge and techniques passed on has seen participants themselves become trainers. Industry organisations are growing capability in bioinformatics as a result. Many of these people have gone on to become Genomics Aotearoa Training Associates.

Not only is this effort helping organisations, but many more students are also able to access free bioinformatics training. Having bioinformatics skills strengthens our future capability as students move into early-career research positions. The viability of bioinformatics as a career option could in turn encourage mathematics as a study option at high school level.

The growth of bioinformatics capability within local research organisations is extending our capacity to use genomics to solve problems in health, the environment and primary production.

Ngoni says, “We have seen our capability growing from a few specialists to a community of genomics people – it started with Dini and me, and now we have an army of experts behind the two of us, enabling us to reach the 1,000-trainee milestone. Our training has developed into a one-stop-shop for those starting projects in genomic data analysis. They ask questions that either we can help with, or we can refer them to people who know the answers.”

All these courses are free, and fully on-line training options have been developed. Attendees range from students to bioinformatics practitioners, managers, and researchers. As well as coming from our partner organisations, we have hosted learners from the University of Canterbury, Lincoln University, University of Melbourne, Auckland University of Technology, Unitec Institute of Technology, National Institute of Water and Atmospheric Research, GNS Science, Livestock Improvement Corporation, Museum of New Zealand Te Papa Tongarewa, Malaghan Institute, Ministry for Primary Industries, Ministry of Health, and the Cawthron Institute.

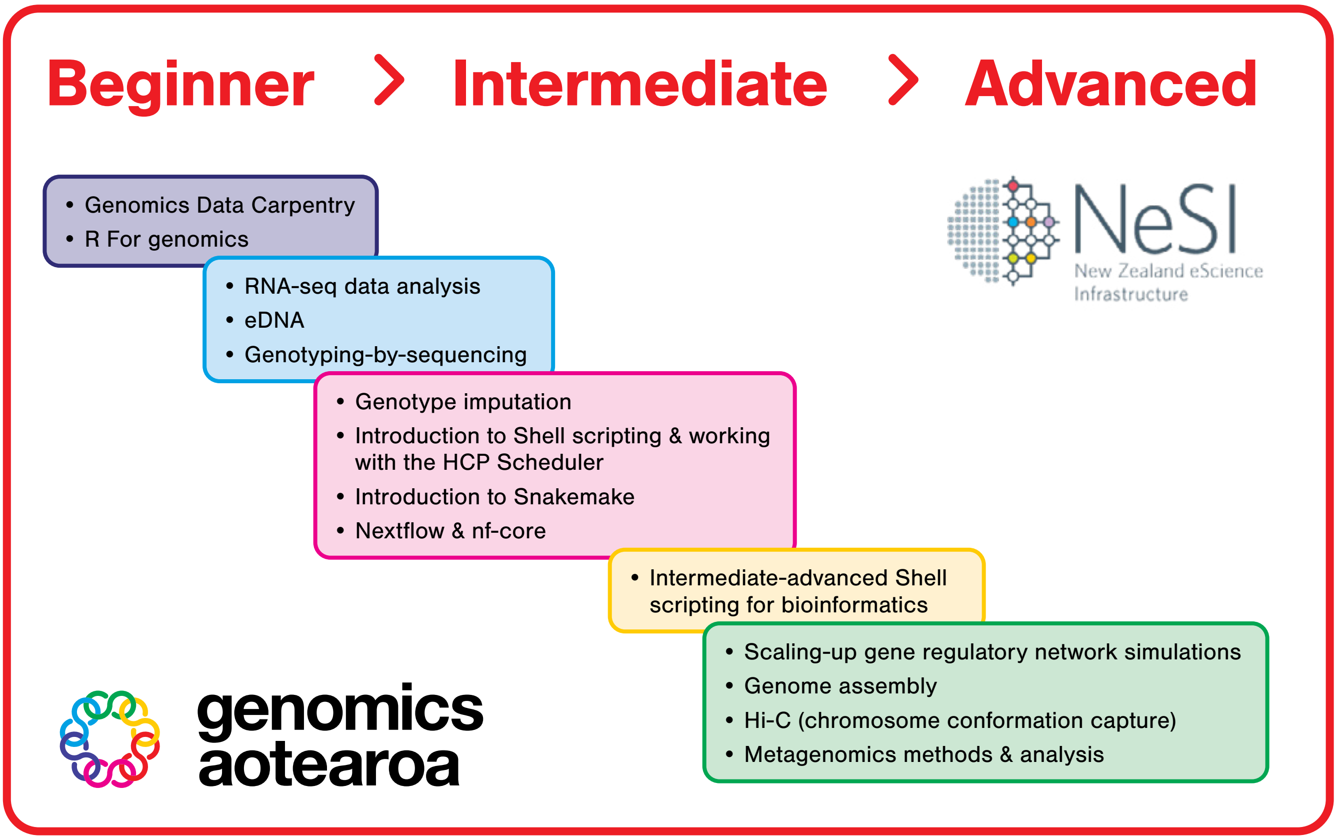
Under the direction of newly appointed Training Coordinator Tyler McInnes the programme will further expand to delivering customised software environments and data management services.

At the same time Bioinformatics Trainer Jian Sheng Boey is helping to organise and share the workflows from bioinformatics researchers via Genomics Aotearoa’s GitHub site, making their pipelines and processes accessible for all.

Another achievement is that the knowledge and techniques passed on has seen participants themselves become trainers. Industry organisations are growing capability in bioinformatics as a result. Many of these people have gone on to become Genomics Aotearoa Training Associates.

Not only is this effort helping organisations, but many more students are also able to access free bioinformatics training. Having bioinformatics skills strengthens our future capability as students move into early-career research positions. The viability of bioinformatics as a career option could in turn encourage mathematics as a study option at high school level.

The growth of bioinformatics capability within local research organisations is extending our capacity to use genomics to solve problems in health, the environment and primary production.



Beginner > Intermediate > Advanced

- Genomics Data Carpentry
- R For genomics

- RNA-seq data analysis
- eDNA
- Genotyping-by-sequencing

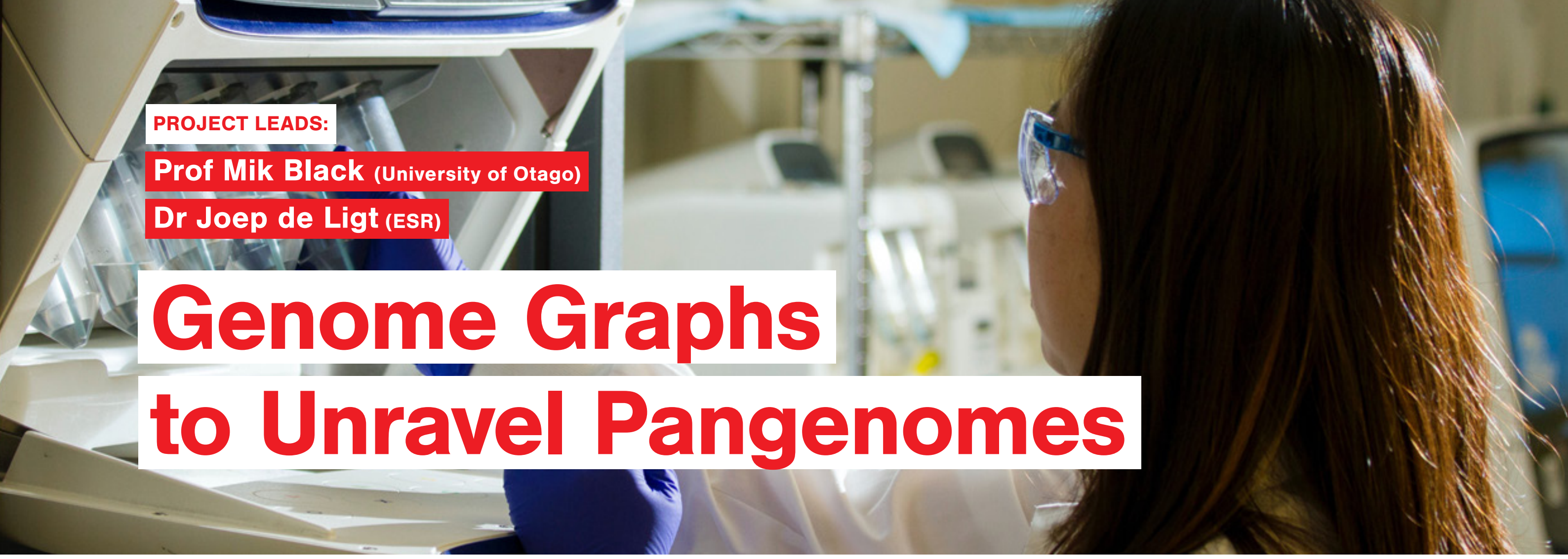
- Genotype imputation
- Introduction to Shell scripting & working with the HCP Scheduler
- Introduction to Snakemake
- Nextflow & nf-core

- Intermediate-advanced Shell scripting for bioinformatics

- Scaling-up gene regulatory network simulations
- Genome assembly
- Hi-C (chromosome conformation capture)
- Metagenomics methods & analysis

NeSI
New Zealand eScience Infrastructure

genomics
aotearoa



PROJECT LEADS:

Prof Mik Black (University of Otago)

Dr Joep de Ligt (ESR)

Genome Graphs to Unravel Pangenomes

Pangenome analysis is increasingly being used to assess the entire genomic variation in a population, or even a species. But pangenome analysis is a challenge for bioinformatics. Graph-based analyses offer an improved approach to detect genomic variation. Although graph-based bioinformatic methods are relatively new, they are experiencing strong uptake internationally, and have been used in several new Genomics Aotearoa projects.

Genome Graph methods are being used to improve variant detection in a collection of case studies. The case studies include both large (human-scale) and small (microbial-scale) genomes that are relevant for Aotearoa. Graph-based methodology is applicable to a range of Genomics Aotearoa projects across all themes, including the High Quality Genomes and Population Genomics, Environmental Microbiomes, Variome and Rakeiora projects. These projects all provide ideal models for case studies.

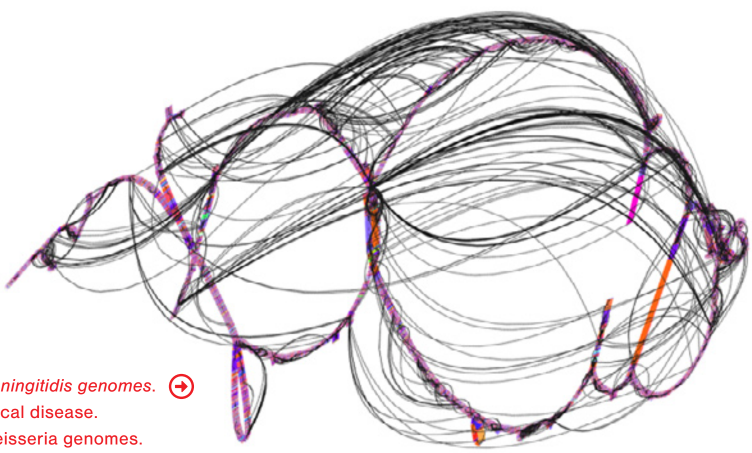
The outputs of the project will include benchmarked tools and workflows for genome graphs for use in blue-sky research, public health, and primary production, as well as training for researchers in Aotearoa on the use of the methods.

Despite the disruptions caused by COVID-19, the team has made a strong start by establishing a key collaboration with one of the world's leading graph-genome research groups, that of Erik Garrison at the University of Tennessee Health Science Center (ekg.github.io/). Genome Graphs project team member Zoe Yang is an active participant in the Garrison research work. The Garrison group benefit from the Genome Graphs project research as it adds organismal diversity to complement their mammal-focussed research.

The collaboration has already resulted in several changes to the software tool that is most versatile and applicable to Aotearoa's graph-genome needs. The first version of a best practice pipeline for use of graph genomes in microbial genomics is now publicly available on Github (github.com/GenomicsAotearoa/GraphPanGenomes_Microbial). This pipeline was also shared through the structural variant (SV) detection workshop organised by Tammy Steeves (Canterbury University). Zoe Yang's presentation triggered a huge interest among participants to apply the pipeline to their own datasets. This interest demonstrates that it was timely for Genomics Aotearoa to have a project focussed on practical uses of graph-genome technologies.

Building on the Genome Graphs project foundations, Zoe Yang has been awarded Strategic Science Investment funding to use graph genome methodologies for plasmid surveillance to track antibiotic-resistant pathogenic bacteria.

Now that the full team is in place, they are looking forward to an exciting year of further progress and will be delivering training sessions on the use of these technologies.



2D visualisation of pangenome graph derived from 24 *Neisseria meningitidis genomes.* *Neisseria meningitidis* is the bacterium responsible for meningococcal disease. The black lines in the image indicate extensive recombination in Neisseria genomes.

Cultural Infrastructure

to Build Te Ao Māori into the Platform

Genomics Aotearoa infrastructure is not just about the technical platform used by researchers. Cultural infrastructure demonstrates Genomics Aotearoa's commitment to a framework that supports Māori participants, whether they are researchers, students, partners or end-users.

> Genomics Aotearoa has always engaged in Māori consultation, but the new Indigenous Genomics Platform takes this much further – these projects have been designed by Māori, and will bring direct benefit to Māori, be that a Māori business or a local community. It will also help build trust in genomic approaches
>
> – **Tracey Godfery**
> Genomics Aotearoa Vision
> Mātauranga Manager

PROJECT LEAD:

Prof Mik Black (University of Otago)

The Aotearoa Genomic Data Repository: A Nationally Significant Database for Taonga Species

The Aotearoa Genomic Data Repository (AGDR) has been implemented in accordance with Te Mata Ira and Te Nohonga Kaitiaki guidelines and adheres to the principles of Māori Data Sovereignty. This repository helps organisations or individuals exercise their kaitiakitanga over taonga, while facilitating controlled data management and sharing.

Hosted at NeSI, the AGDR stores genomic data from taonga. Taonga include plant and animal species, but the definition can also extend to environmental samples and even bioactive compounds derived from plants and animals. Local hosting is a critical requirement because if data are outside Aotearoa’s jurisdiction, Te Tiriti o Waitangi rights become extremely difficult to enforce.

Tikanga-based data submission and retrieval ensures data rights are correctly assigned. Submitters must provide information about the consultation process that was followed and the kaitiaki of the data being submitted. This information is also used for data access: requests must be approved by the kaitiaki of that data set.

Initiated in 2019, the AGDR has been progressively enhanced over the last three years. The ability to incorporate mātauranga and biocultural labels has recently been added to the AGDR. A Māori advisory group (Ngā Kaiārahi) is being organised to provide ongoing cultural guidance on the operation of the data repository, and on other activities within the Bioinformatics 2 project.

A major goal achieved over the past year was to move the AGDR to a full production system. New data sets are being deposited, and requests for access are being passed on to kaitiaki. The AGDR now hosts genomic data from 12 species and an eDNA dataset.

CASE STUDY:

Data sovereignty in action

Considerable consultation and work went into developing the AGDR repository, both from a technical perspective and to implement Māori Data Sovereignty principles to meet Te Tiriti o Waitangi obligations. This is the first time a comprehensive and Aotearoa-wide set of protocols using CARE principles[1] has been applied to taonga data.

Project co-lead Mik Black says, “We wanted to deploy a system that allows both researchers and kaitiaki to fulfil their responsibilities relating to genomic data sets. We have done extensive work that provides a strong foundation - we now have a robust facility with systems to move data in and out rapidly using latest technology, using FAIR[2] and CARE principles, have the Te Mata Ira and Te Nohonga Kaitiaki guidelines as an ethical framework for kaitiaki-centric data access and submission protocols based on principles of Māori data sovereignty.”

Approval protocols have now been put to the test. There have been nine applications for access to taonga data over the last year. Three of these applications were approved by kaitiaki, two were declined, and four require more information before a decision is made.

The two declined applications were both missing basic information about the projects. Enabling kaitiaki to make informed decisions on the use of data and rejecting applications is an example of true data sovereignty.

Aotearoa has started the process of keeping taonga genomic data safe from exploitation by housing data in a locally based platform and involving kaitiaki in decisions about data access. Although the AGDR cannot rectify historical misuse of data, the AGDR data sovereignty process has now been tested and found robust. The way is paved for kaitiaki to have the final say on of the use of taonga data for the long-term benefit of Aotearoa.

---

[1] **C**ollective benefit, **A**uthority to control, **R**esponsibility and **E**thics

[2] **F**indable, **A**ccessible, **I**nteroperable, **R**eusable

PROJECT LEAD:

Prof Mik Black (University of Otago)

Te Nohonga Kaitiaki: Guidelines for Genomic Research on Taonga

The Te Nohonga Kaitiaki (Place of the Guardians) guidelines are designed to ensure the planning and execution of genomic research on taonga incorporates tikanga and respects kaitiaki, and mātauranga Māori. Te Nohonga Kaitiaki guidelines' primary audience is the science community, to help guide them when researching taonga and engaging with iwi/hapū/whānau. The guidelines also help Māori groups design protocols for research relevant to their own tikanga and mātauranga. The roles of kaitiaki of taonga and the responsibilities of institutional stewards are explicit in the guidelines.

The Te Nohonga Kaitiaki project was completed last year. The guidelines have been published and are available in digital form on the Genomics Aotearoa website (genomics-aotearoa.org.nz/resources/guidelines-maori-genomics-research), and printed copies are available from the researchers. Maui Hudson provides an introduction that describes the purpose and cultural basis for the guidelines on YouTube (youtube.com/watch?v=axe6STfTxe4).

Genomics Aotearoa strongly encourages all researchers to read and absorb the guidelines to gain a better understanding of genomics research on taonga. While taonga include native species themselves, introduced species can also be taonga depending on their relationship with people. Microorganisms from environmental samples and bioactive compounds derived from species can also be taonga. Therefore, it is critical that all researchers working on environmental topics explore the relevance of the guidelines at the outset of research planning.

Outside of work on taonga, important guidelines to be consulted when considering research projects involving people include:

- Te Mata Ira: Guidelines for genomic research with Māori (genomics-aotearoa.org.nz/sites/default/files/2019-03/Te-Mata-Ira-Genome-Research-Guidelines.pdf).
- He Tangata Kei Tua: Guidelines for biobanking with Māori (genomics-aotearoa.org.nz/sites/default/files/2019-03/He-Tangata-Kei-Tua-Biobanking-Guidelines.pdf).
- Te Ara Tika: Guidelines for Māori Research Ethics (genomics-aotearoa.org.nz/sites/default/files/2021-11/Te Ara Tika Guidelines for Māori Research Ethics.pdf).

CASE STUDY:

Te Nohonga Kaitiaki guidelines define a framework for engagement

Respecting kaitiaki interests over taonga demonstrates a commitment to Te Tiriti o Waitangi. Until now, there has been little guidance to ensure that taonga (any living thing, or their derivatives, of value to Māori) are being studied in a way that upholds Te Tiriti o Waitangi.

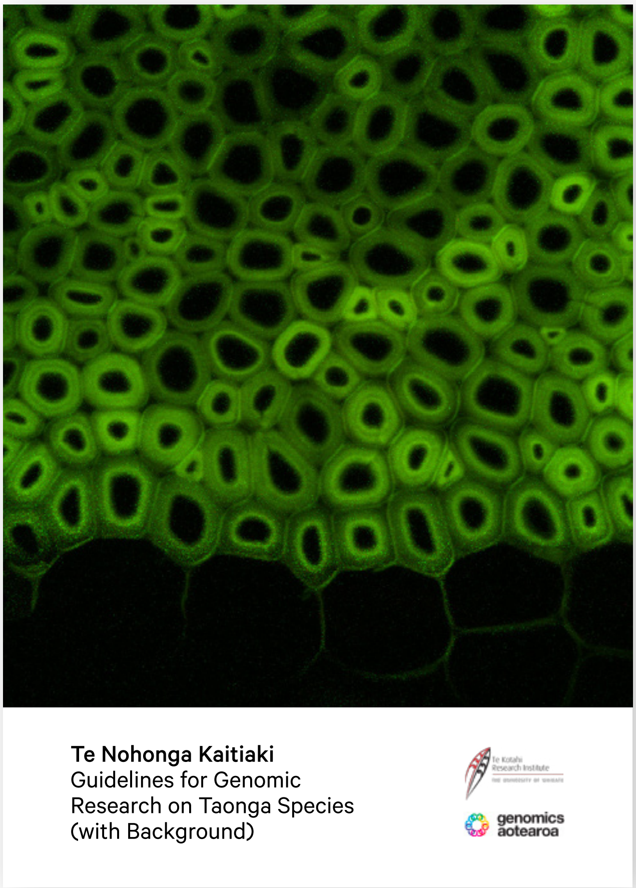
All research involving genetic resources for the purposes of conservation and ecology within Aotearoa is of value and interest to kaitiaki.

Genomics Aotearoa researchers recognised the need for a comprehensive framework for research positioned at the intersection of genomics, innovation, and Te Ao Māori.

Consultation to develop the framework involved two rounds of nationwide hui, presentations and wānanga attended by science and community representatives, government departments and Crown entities, together with submissions from people involved in the consultation.

The outcomes from this engagement evolved into Te Nohonga Kaitiaki, which contains several important elements:

- An Engagement Framework, which illustrates considerations for effective engagement at different levels of a project, from inception to completion.
- Guiding Principles, which set waypoints to frame the thinking around genomic research.
- Operating Principles, which reflect the nature and relationship whānau, hapū and iwi hold with their respective taonga. The Operating Principles provide clarity around how relationships with taonga also have relationships with people and place.
- An Engagement Checklist provides questions to inform the development of robust relationships with Māori. An important aspect of conducting genomic research that involves taonga is the need to engage with the right people in the design phase.



Te Nohonga Kaitiaki
Guidelines for Genomic
Research on Taonga Species
(with Background)

Te Kotahi Research Institute

genomics aotearoa

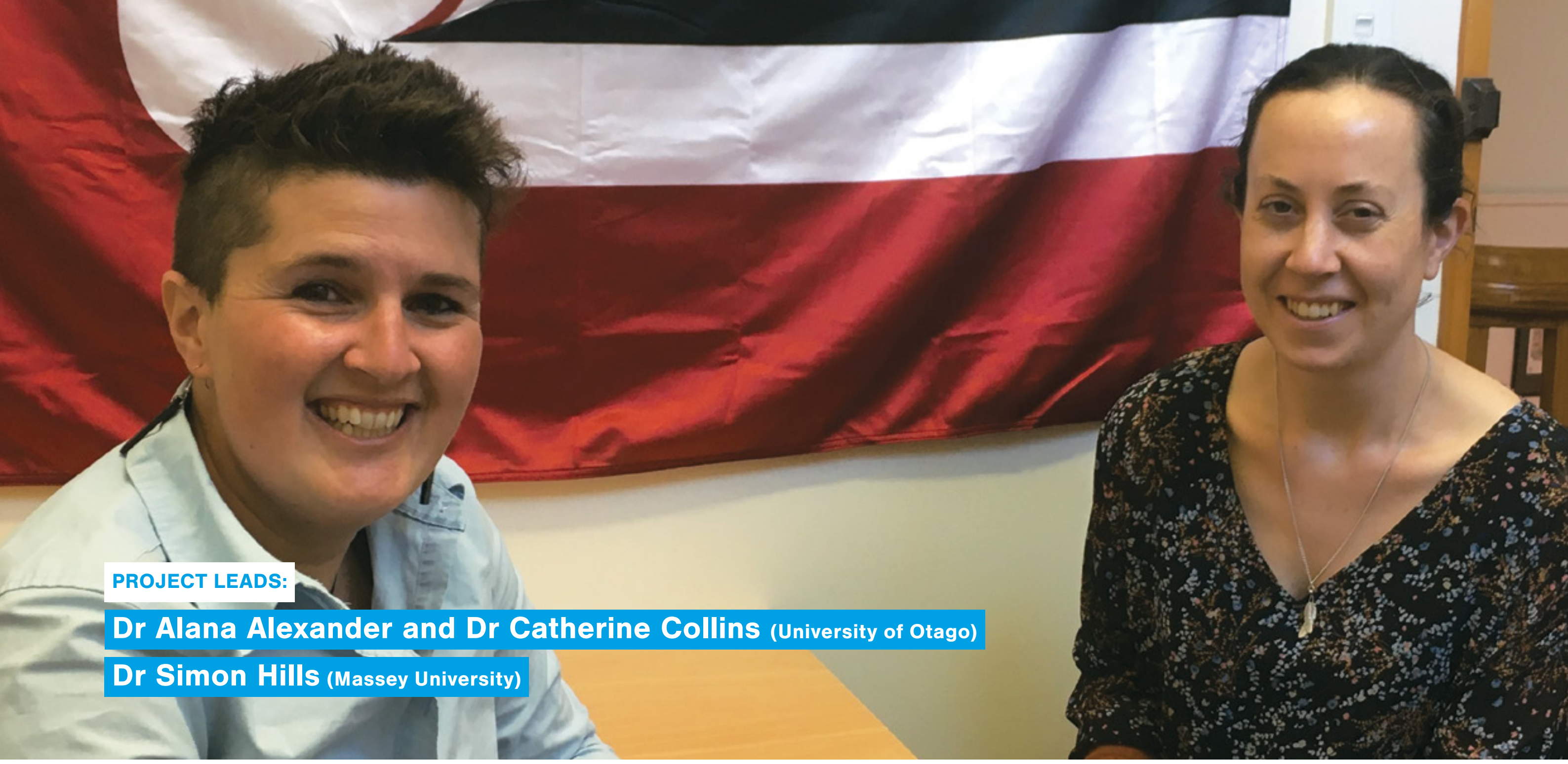
All researchers working in Aotearoa can now access a framework that guides effective engagement with Māori across three responsiveness levels – project, organisation, and systems.

Since completion the guidelines have been shared with the science research community around Aotearoa and are used in undergraduate and graduate teaching at the University of Otago. More than 1,000 printed copies of the guidelines have been distributed amongst key institutions.

Discussions have also started around the potential transferability of the guidelines to other academic fields, such as archaeology. The Australian Institute of Marine Science has shown interest in how the guidelines provide an example of ways to support their work with Indigenous peoples. The guidelines were well-received by Plant & Food Research at a wānanga on research on taonga. Feedback on the guidelines has been positive.

Genomics Aotearoa now requires all new projects in the environmental research space to consult the Engagement Checklist at a minimum during their design phase.

Project Level Responsiveness	Organisastion Level Responsiveness	System Level Responsiveness
Project Outcomes	Capacity Building	End Users
Level of Involvement	Embedding Relationships	Research Networks and Consortia
Intellectual Contribution of Māori/Mana Whenua	Sample/Data Access and Governance	International Agreements
Engagement/ Communication	Benefit Sharing	Research Funding



PROJECT LEADS:

Dr Alana Alexander and Dr Catherine Collins (University of Otago)

Dr Simon Hills (Massey University)

Ruatau: Connecting Māori Genomic Scientists and Communities

Empowering Māori communities to make informed decisions about important genomics issues such as hauora (health) of people, or kaitiakitanga (guardianship) of taonga is vital. Māori scientists (kaipūtaiao Māori) can help through understanding community perspectives, priorities, and concerns. But Māori are under-represented in academia and in genomics research. Furthermore, many Māori genomics researchers grew up Te Ao Pākehā and feel that they need more grounding in Te Ao Māori.

Led by Māori early-career researchers, Ruatau seeks to develop processes and resources to increase the capability of both Māori communities and Māori scientists to achieve and contribute to genomic aspirations. Specifically, Ruatau supports an increase in unique whānau-led research, culturally informed genomics research, Māori participation in research, as well as the creation of a network of Māori with interests in genomics.

At the end of the first year of Ruatau, the researchers are in the resource-building phase for this project. Some of these resource materials are now available for end-users to access. The most exciting outcome of Ruatau over the past year was a kanohi ki te kanohi hui of the project leads and advisors in May. This hui allowed the team to refine the direction of the project and get feedback on resources developed to date.

The researchers have identified the hapori Māori to initially approach and are in the process of generating resources for culturally upskilling kaipūtaiao Māori who grew up Te Ao Pākehā, for genomics wānanga, and for the literature review that forms the academic foundations for Ruatau.

The project explicitly focuses on Māori researchers and communities and is being implemented following kaupapa Māori best practices for engagement, including co-design. All project leads and advisors are Māori, developing these researchers, as well as supporting the development of leaders in genomics from local communities. As co-lead Alana Alexander says, “We are really excited to be able to bridge being Māori and being genomic scientists, while hopefully helping the hapori Māori we work with to realise their genomics aspirations.”

The coming year will see an initial genomics wānanga and engagement of a postdoctoral researcher who will lead the delivery of an efficacy wānanga and analysis of resulting data. This process will help researchers evaluate the benefits and room for improvement of Ruatau.

PROJECT LEAD:

Assoc Prof Maui Hudson (University of Waikato)

SING Aotearoa: Mentoring Emerging Māori Genomics Researchers

Genomics research is still an emerging discipline that has significant potential benefits for Māori, particularly in the environment, health, and socio-economic sectors. The Summer Internship for Indigenous Genomics (SING Aotearoa) initiative delivers a training programme for pakeke and tauira Māori interested in learning about genomics and bioinformatics, and the ethical and cultural issues around genomics research.

SING Aotearoa brings tauira Māori interns together with genomics experts in an intense five-day marae-based (usually) residential programme. The programme also provides opportunities for non-Māori researchers to facilitate sessions and engage with participants. The initiative creates internships, research collaborations, and scholarship pathways.

SING Aotearoa maintains a strong relationship with the wider SING Consortium (singconsortium.org) and continues to create opportunities for faculty and intern exchanges. The SING Consortium is collaborating with ENRICH (enrich-hub.org) to engage in international discussions around digital sequence information and benefit-sharing and the potential for other collaborative activities.

This year SING 2022 was originally to be held in Te Tai Tokerau. However, due to COVID-19 concerns, the programme was delivered online through Zoom over three days in June 2022. A benefit of the online format was that it allowed alumni from previous years to also attend. The new interns' first experiences of SING were buoyed by these alumni re-engaging with the programme. The presence of a kaumatua also added an exciting new dimension, and the SING committee plans to discuss a permanent kaumatua role for SING Aotearoa.

Although Zoom fatigue and a preference for kanohi ki te kanohi wānanga were challenges to SING 2022, 23 attendees took part. This level of attendance demonstrates that despite the challenges, there is still enthusiasm to participate. As with previous SING workshops, all presenters and faculty donated their time.

This year the interns were highly positive about their SING experiences. Their comments included:

> *I gained a better understanding of the cultural perspectives, and the struggles Indigenous scientists and academics are facing and continue to face.*
>
> *SING highlights the importance of connection with other indigenous academics and creating a community in spaces that may be daunting as an indigenous person.*
>
> *Most valuable to me was talking about the value of tautoko, mentors, and community as indigenous researchers, and how to navigate predominantly non-indigenous spaces more safely.*
>
> *I gained a stronger understanding of how we can be leaders in this space and support each other and those coming through (also the power of collaboration).*
>
> *"Turn around and look behind you." Where you are now is where many dream to be, so do not underestimate your value and experience.*

A benefit of online delivery is lower cost. Consequently, the surplus from SING 2022 will enable the team to offer enhanced opportunities to SING alumni and an increase to intern numbers for SING 2023, which will be held in Te Tai Tokerau.

SING Aotearoa continues to draw on expertise from throughout the country and seeks contributions from genomic researchers involved in Genomics Aotearoa funded projects, thereby creating additional exposure and dual capability building for both our interns and the facilitators.

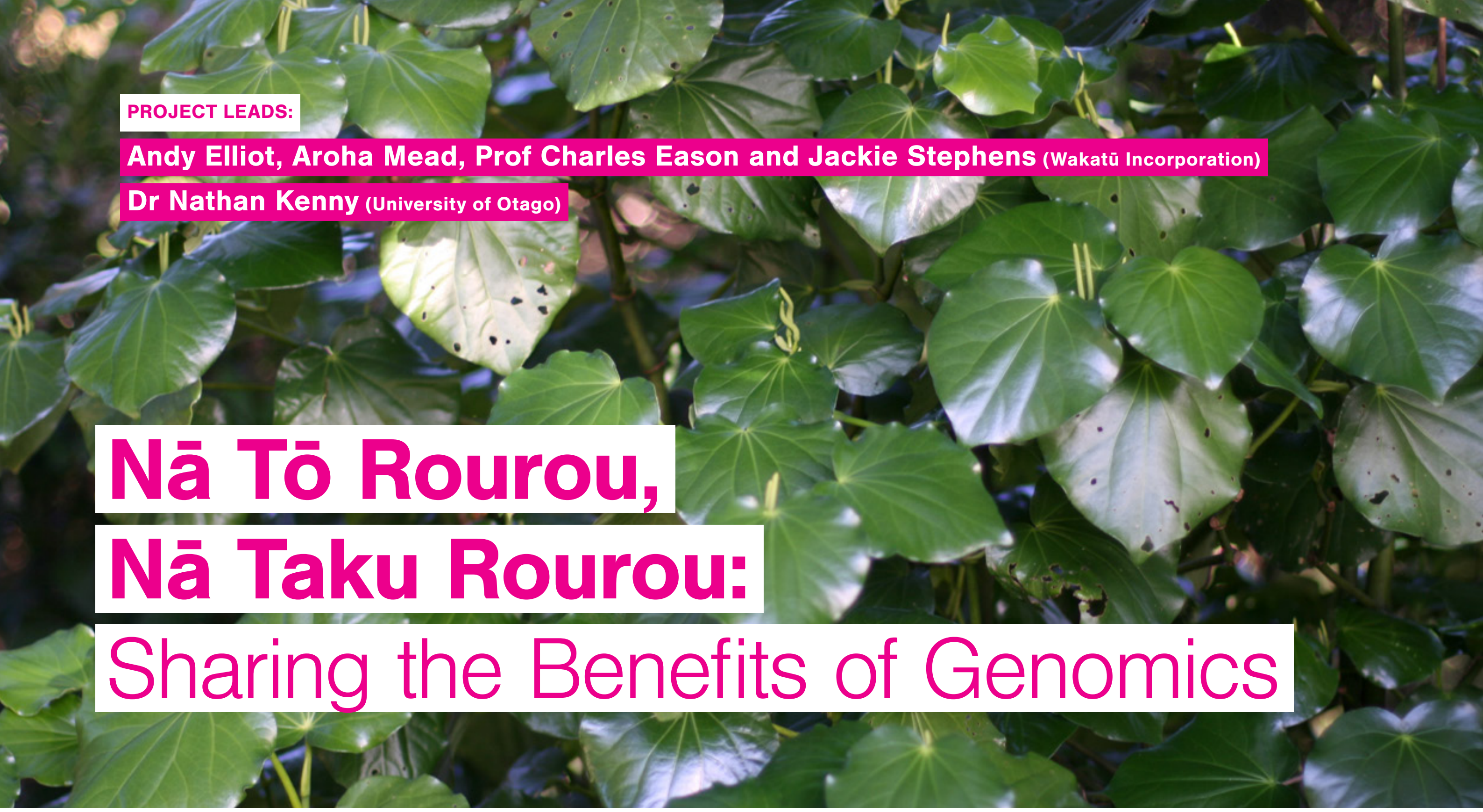
The SING Aotearoa committee is also pleased to have the continued support of the Maurice Wilkins Centre, which for the next seven years will co-fund scholarship opportunities for SING alumni to attend Queenstown Research Week – six alumni are attending later in 2022.

SING remains a key capability-building initiative supported by Genomics Aotearoa.

Upcoming Projects

Genomics Aotearoa's project selection criteria have always been consistent with the overall strategic direction of the platform. New this year, the Indigenous Genomics Platform (IGP) increases the ability of Māori groups to lead projects and engage directly with genomics researchers, and for our research to better reflect Māori end-user needs.

Like our other projects, proposals for the IGP go through a rigorous selection process, with the additional requirements of Māori-led research and clear benefits to Māori. The standard of proposals received has been high, and three IGP projects have so far been selected to start in the coming year.



PROJECT LEADS:

Andy Elliot, Aroha Mead, Prof Charles Eason and Jackie Stephens (Wakatū Incorporation)

Dr Nathan Kenny (University of Otago)

Nā Tō Rourou, Nā Taku Rourou:
Sharing the Benefits of Genomics

Nā tō rourou, nā taku rourou aims to better understand the genomic basis of productivity traits of high-value marine species and ensure the ownership of this research is assigned to the right people.

Many initiatives propose to sequence the genomes of all species on earth. In some cases, these initiatives are not sensitive to the Wai 262 Te Tiriti o Waitangi claim or Convention on Biological Diversity and Nagoya Protocol implications. Nā tō rourou, nā taku rourou establishes a model for collaboration among Māori organisations and research institutions, with ownership and governance of genomic data by Māori, using exemplar datasets to test the model.

The project focuses on how the commercial benefits of genomic work can be vested long-term in the kaitiaki of the original sample. In an increasingly globalised world, where genomic sequencing is becoming more routine, we run the danger of losing control of the whakapapa of our taonga.

Wakatū Incorporation's strategy of cataloguing, characterising, and developing a deep understanding of the endemic species of Te Tau Ihu is the perfect vehicle for testing this Māori-centred model for genomic work.

The project will sequence the genomes of kopakopa (the ribbed mussel *Aulacomya atra maoriana*) and karengo (the seaweed *Porphyra* sp.), marine species of significant interest to Wakatū and the iwi of Te Tau Ihu. Genomes of previously identified individuals with traits related to growth rate, adult condition and divergent phenotype will then be re-sequenced, followed by genotype imputation when necessary. The project will generate high density linkage maps, relationship matrices, gene predictions, and other useful resources for breeding and production pipelines.

The team will use the project outputs to test linking genomic data to the rohe from where they originated and consider how the long-term benefits of this work can be tied to mana whenua. The project will prove how genomic information can be gathered, used, and safeguarded by indigenous organisations to drive forward their aspirations. The outputs of genomic sequencing will extend on Wakatū's existing benefit sharing programmes.

The benefits of this genomics-led approach will be relayed to the community through the Te Tau Ihu o Te Waka-a-Maui fisheries forum and other avenues, in line with Wakatū's ongoing commitments to benefit sharing and education.

Anticipated outcomes of the project include:

- A new model for access and benefit sharing for Māori and research organisations.
- Rangatahi trained in novel skills, both in commercial aquaculture and in genetic sequencing technologies.
- Opportunities for both education and employment for Rangatahi.
- Direct commercial benefit to Wakatū Incorporation and its constituent whānau.
- Principles for genetic characterisation of marine organisms, and information that can be used for *blue-sky* science (with the consent of Wakatū).

PROJECT LEADS:

Chris Insley (Te Arawa Fisheries, Smart Māori Aquaculture)

Assoc Prof Peter Ritchie (Te Herenga Waka-Victoria University of Wellington)

Dr Maren Wellenreuther and Dr David Chagné (Plant & Food Research)

Development of Genomic Resources for the Selective Breeding of Kingfish/Haku

Globally and nationally, the demand for seafood is growing. Aquaculture programmes in Aotearoa rely almost exclusively on the farming of three species (greenshell mussels, pacific oysters, and chinook/king salmon), highlighting a need to increase resilience of the sector through diversification.

The Smart Māori Aquaculture initiative (smartmaoriaquaculture.co.nz) is a partnership led by Ngā Iwi i te Rohe o Te Waiariki, together with the Ministry for Primary Industries (MPI) and Te Ohu Kaimoana. The partnership explores pathways to a sustainable, resilient, and world-class Māori aquaculture industry in the Bay of Plenty. Yellowtail kingfish (warehenga or haku) was identified by the initiative as the native finfish with the highest aquaculture potential. To support aquaculture development of kingfish, resources are needed to underpin selective breeding.

Kingfish is an important traditional food for Māori, but no quantitative data on Māori customary non-commercial catch are available. Because of the coastal distribution of the species and its inclination to strike lures, it is likely that historically Māori caught considerable numbers of kingfish.

The project has three key objectives, which are to:

1. Generate a high-quality genome assembly for kingfish.
2. Construct and evaluate the genomic structure and connectivity of wild kingfish populations around Aotearoa.
3. Develop a data and Intellectual Property management framework as a model for domestication of a taonga species in partnership between kaitiaki and researchers.

Anticipated outcomes of the project include:

- Iwi partnerships with academic and research institutes to provide Māori career opportunities in aquaculture, and drive success.
- Māori economic development – growing people's skills through job creation, training, career pathways, and research and leadership opportunities.
- Strengthening the traditional iwi and community relationship with the ocean.
- Empowering and exercising kaitiakitanga and maintaining and enhancing the mauri of Te Moana Nui-ā-Toi.

PROJECT LEADS:

Dr Keith Funnell (Plant & Food Research)

Colan Balkwill (Te Herenga Waka-Victoria University of Wellington)

Realising the Potential of Genomics for Swamp Maire Replanting in the Rangitāne Rohe

The swamp maire (*Syzygium maire*), a nationally critically threatened tree, is the subject of active restoration in the Manawatū region, especially along the new Te Ahu a Turanga (Manawatū Tararua highway). The population genetics of the plant are being studied, but iwi need greater capability and capacity in the use of genomics, and the mātauranga of swamp maire, to enable implementation of genomics knowledge to guide forest restoration.

As iwi and other Māori groups across Aotearoa reaffirm and expand their work as kaitiaki of taonga species, there is growing demand for skills and data that supports Māori aspirations. This project enables the participation of Rangitāne o Manawatū in the Aotearoa science system, facilitated by researchers at Plant & Food Research and Te Herenga Waka-Victoria University of Wellington. Rangitāne o Manawatū will gain access to a broad range of scientific expertise to aid in developing innovative, mātauranga Māori-led kaitiakitanga of swamp maire, and implementing genomics-based solutions to facilitate sustainable forest ecosystems. In return, Rangitāne o Manawatū will support scientists in gaining an authentic understanding of mātauranga and kaitiakitanga.

Populations of the swamp maire are small and fragmented, typically surrounded by farmland. Where remnants do exist, they are often unable to regenerate due to drainage of habitat and, more recently, the arrival of myrtle rust in Aotearoa, a significant fungal pathogen. With fewer than 12 mature trees of swamp maire within the rohe, the lack of genetic diversity may hinder the establishment of a naturally sustainable forest population through the replanting programme alone.

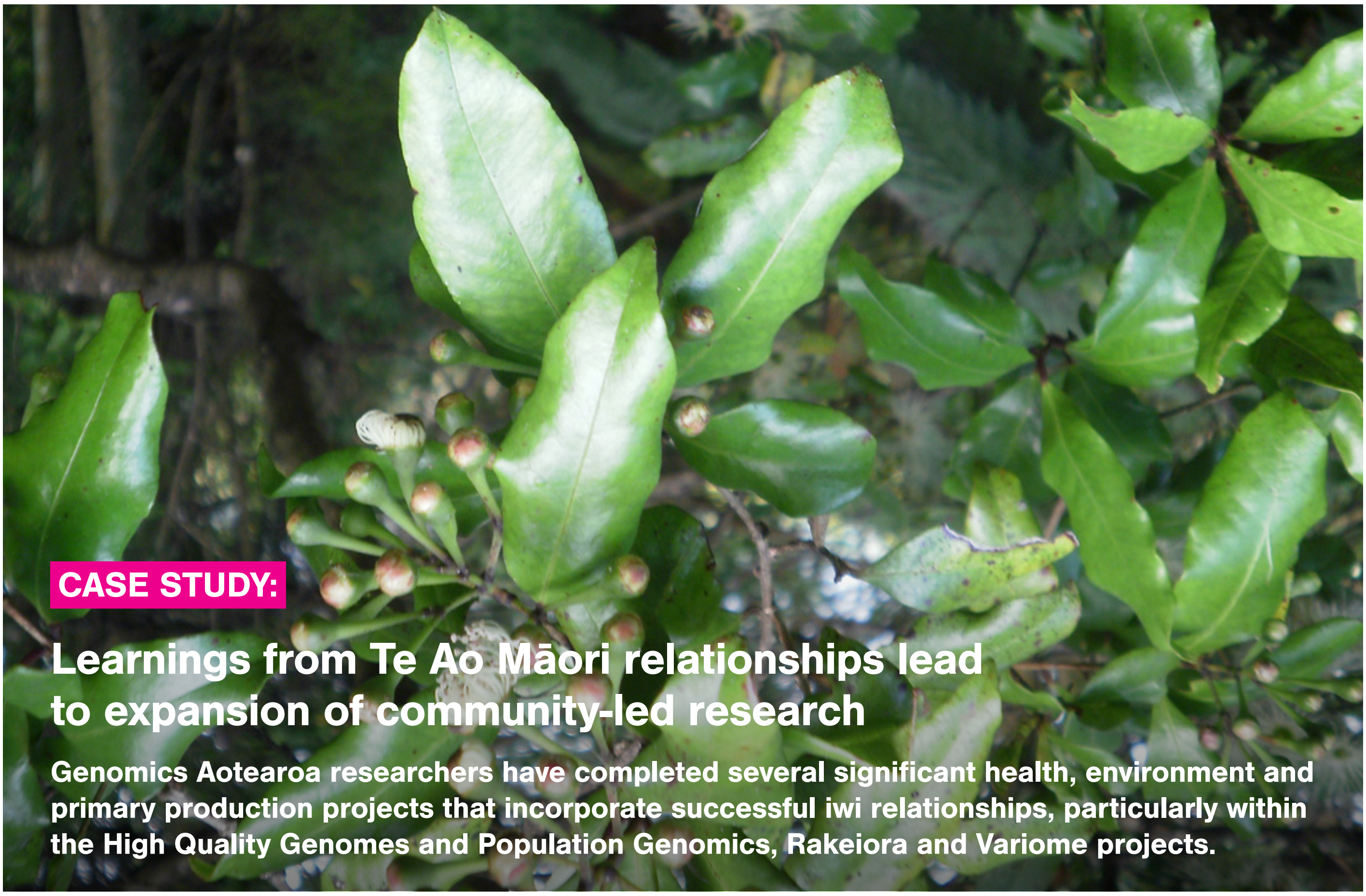
Genomic knowledge will enhance conservation and restoration strategies through prioritisation of populations for protection, seed-preservation, propagule sourcing and the development of restoration strategies to facilitate genetic adaptation.

The project will increase understanding of the species genetic diversity, its capacity for adaptation, and inform the ability to select and propagate plant material. Together with habitat suitability mapping, this information will be built into an integrated restoration model for the Manawatū, allowing finer scale identification of suitable planting locations and genetic material within the rohe.

Protocols for implementing the genomics knowledge into the restoration of a plant species will be developed by Rangitāne o Manawatū as a model, with the creation of a video to document this journey, allowing others to visualise and share this experience.

Anticipated outcomes of the project include:

- Enhanced Māori capability to implement genomics solutions for restoration of a threatened (nationally critical) long-lived plant species.
- Iwi partnerships with academic and research institutes that offer powerful opportunity to grow Māori careers in plant conservation, and drive success.
- Strengthening traditional iwi and community relationships with taonga.
- Empowering and exercising kaitiakitanga and maintaining and enhancing the mauri of swamp maire.



CASE STUDY:

Learnings from Te Ao Māori relationships lead to expansion of community-led research

Genomics Aotearoa researchers have completed several significant health, environment and primary production projects that incorporate successful iwi relationships, particularly within the High Quality Genomes and Population Genomics, Rakeiora and Variome projects.

The challenge is to extend the scope from individual researchers engaging with specific iwi or hapū, into a wider and more inclusive model that is more responsive to Māori needs, and that ultimately builds capability and capacity for Māori-led research.

Genomics Aotearoa leadership needed strengthening at a governance level to better provide guidance, assist with relationship building and extend its reach into the Māori community. Disestablishing its Kāhui and appointing three Māori board members – representing health, the environment and mana whenua has given equal representation to Māori, at the highest level.

As well as new board representation, three new Māori members have been appointed to our Science Leadership Team, and our Vision Mātauranga Manager Tracey Godfery completes Genomics Aotearoa's new Māori leadership.

Led by a Te Ao Māori Strategy, the Indigenous Genomics Platform (IGP) ensures projects are community-driven rather than researcher-focussed. The IGP uses a dedicated $2.4 million fund to target support for projects that demonstrate partnership and provide clear benefit to Māori.

"The new Indigenous Genomics Platform presents valuable opportunities to engage more cohesively and form enduring Māori partnerships, extending the iwi relationships already formed by individual researchers working in Genomics Aotearoa's health, environment and primary production projects," says Tracey.

The key criteria developed for funding Māori co-led and co-designed projects through this platform are:

- Benefit and contribution to Māori Community.
- Level of Māori Involvement.
- Clear and effective engagement pathways with Māori.
- Ground-up approach that builds relationships and trust.
- Demonstrated need and clear rationale for the research.
- Feasibility of achieving outcomes within the timeframe.

The IGP fund opened for submissions in January 2022; and the Genomics Aotearoa Vision Mātauranga team reviewed an initial 12 submissions. Three Māori co-led, and co-designed projects were initially selected, with more project awards in the pipeline.

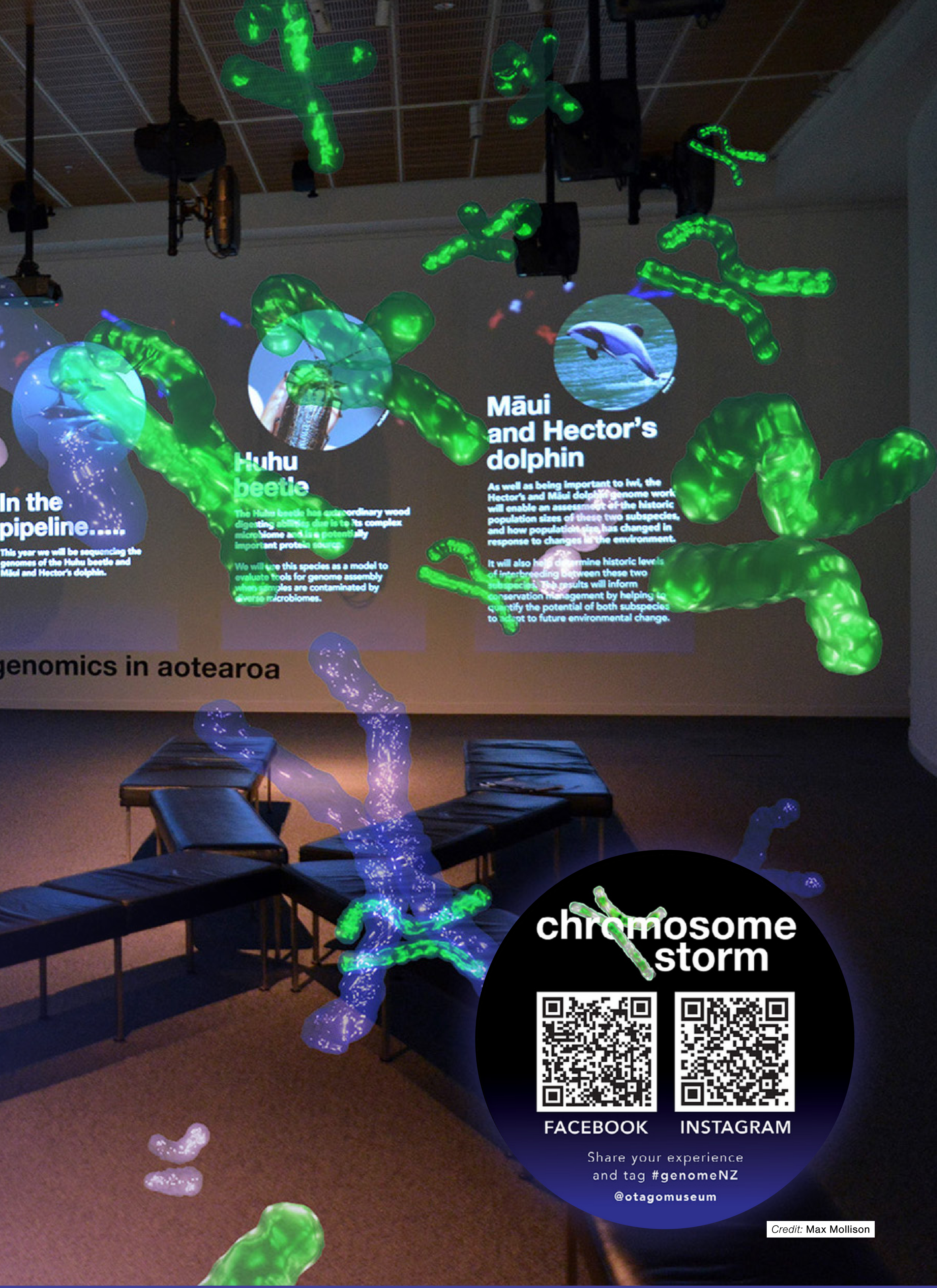
The IGP promotes a consolidated approach to research and engagement, supporting collaborative research that draws on the skills of our genomics researchers as well as the knowledge and skills of the iwi/hapū involved. This approach contributes to successful outcomes for not only the research participants but for Aotearoa as a whole.

Extending Genomics Aotearoa's current portfolio of projects in partnership with Māori promises exciting new opportunities to develop capability and capacity, ensuring more Māori can lead and participate in genomics research in the future.

The IGP is a major collaborative effort that has created considerable interest from organisations interested in strengthening research partnerships with Māori. We expect the IGP will contribute to creating an enduring foundation for Māori engagement and leadership within genomics.

Public Engagement

Throughout the year Genomics Aotearoa has continued a fortnightly Friday seminar series, interactions with the media, monthly newsletter, and public engagement through social media via our YouTube channel and Twitter. Our major public event for the year was the genomics exhibition at Tūhura Otago Museum.



In the pipeline.....

This year we will be sequencing the genomes of the Huhu beetle and Māui and Hector's dolphin.

Huhu beetle

The Huhu beetle has extraordinary wood digesting abilities due is to its complex microbiome and is a potentially important protein source.

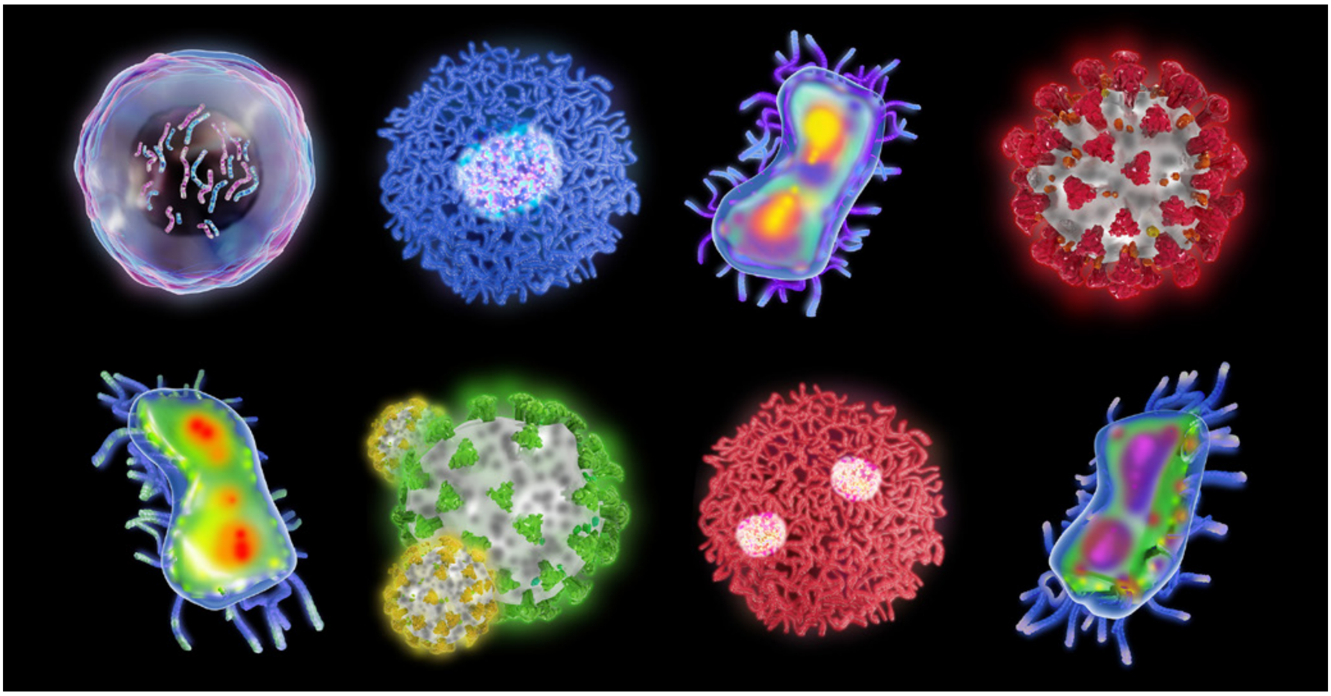
We will use this species as a model to evaluate tools for genome assembly when samples are contaminated by diverse microbiomes.

Māui and Hector's dolphin

As well as being important to iwi, the Hector's and Māui dolphin genome work will enable an assessment of the historic population sizes of these two subspecies, and how population size has changed in response to changes in the environment.

It will also help determine historic levels of interbreeding between these two subspecies. The results will inform conservation management by helping to quantify the potential of both subspecies to adapt to future environmental change.

Credit: Max Mollison



Genome: Science of Life

The major event for us this year was the Genome: Science of Life exhibition, in partnership with Tūhura Otago Museum. The exhibition highlighted the power of genomics research and de-mystified the technology for the public

"Tūhura Otago Museum has the largest science centre and science outreach team in the country," says Marketing Manager Kate Oktay. "Our mandate is to inspire curiosity, so we were thrilled to partner with the Genomics Aotearoa team on this, creating a visual aspect to communicate the science," she continued.

Some of the more remarkable uses of genomics featured in the exhibition included how long-term space travel might affect astronauts, how DNA can both help and hinder criminal investigations, and how genomics can increase our understanding of the mechanisms of disease transmission.

Real time sequencing by Nathan Kenny (Ngāi Tahu and Te Ātiawa), Louise Walker and Katerina Achilleos from the University of Otago was a highlight, with 130 tamariki taking part.

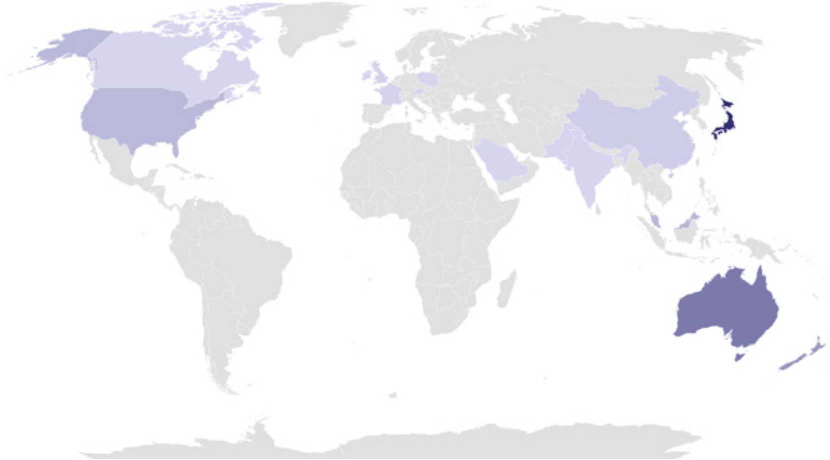
The museum recorded 4,212 visitors and the social media campaign was viewed 97,485 times with a reach of 55,904 people. The museum team say this is an amazing success for an exhibition. The success of the exhibition means it will be held again from November this year to February 2023.

Asia-Pacific Genetics Seminar Series

Genomics Aotearoa is the host of the newly established Asia-Pacific Genetics Seminar series, a collaboration between the Genetics Society of Japan (GSJ) and the Genetics Society of Australasia (GSA). The series, delivered via Zoom, aims to further develop the field of genetics, and promote international exchange among researchers in Asia and Oceania.

The first speaker was Professor Emerita Dr Tomoko Ohta, a ground-breaking world authority on population genetics. Professor Ohta's seminal *nearly neutral theory of molecular evolution* provides a framework to understand the evolution of complex systems.

The talk was extremely well-attended, with 259 participants from 15 countries, as shown in the map below.



Media Engagement

Engagement with media this year included releases on the completion of the world's first Bilberry genome with Plant & Food Research, and a new tactical genetic pest control project on rats with Predator Free 2050.

The Predator Free 2050 project, which explores potential genomic control methods for difficult to eradicate invasive species, has attracted considerable interest locally and internationally. The Discovery Channel plan to run a feature on the project in the coming year.

eResearch NZ 2022

The theme for eResearch NZ 2022 was *Building Capability Together|Waihanga Āheitanga Kotahitanga*., which is very much in line with Genomics Aotearoa's goals. We again co-hosted the conference alongside NeSI and REANNZ in partnership with Te Whare Wānanga o Waitaha University of Canterbury. The event ran from February 9-11, 2022, and for the first time attendance was fully online.

The event is a very effective way of uniting the bioinformatics community in Aotearoa, and this year attracted 301 local and international registrations, including 113 students, and 73 speakers. The Rakeiora project team gave a keynote presentation, and Genomics Aotearoa researchers contributed to the *Te Ao Māori in eResearch* panel

Friday Seminar Series

We have hosted 20 Friday seminars on Zoom over the reporting period, which continue to attract interest from researchers in Aotearoa and beyond. An average of 74 people attended each fortnight, and four seminars attracted more than 100 people.

The Zoom format is useful for both attracting international presenters and for recording the seminars and sharing them via our YouTube channel. This year we had two international speakers, Sook Jung and Alice Roberts. We have four international speakers planned within the next six months.

YouTube Videos

Our YouTube channel (youtube.com/c/GenomicsAotearoa/videos) is home to recordings of our seminars and other videos of interest to the genomics community in Aotearoa. Currently 21 videos are available, and the year ended with 105 subscribers (and counting). Our most popular recording, with more than 8,000 views is by Professor Alice Roberts. Alice is biological anthropologist, biologist, television presenter and author at the University of Birmingham, England. Her seminar is titled *Ancestors: A History of Britain Through Burials.*

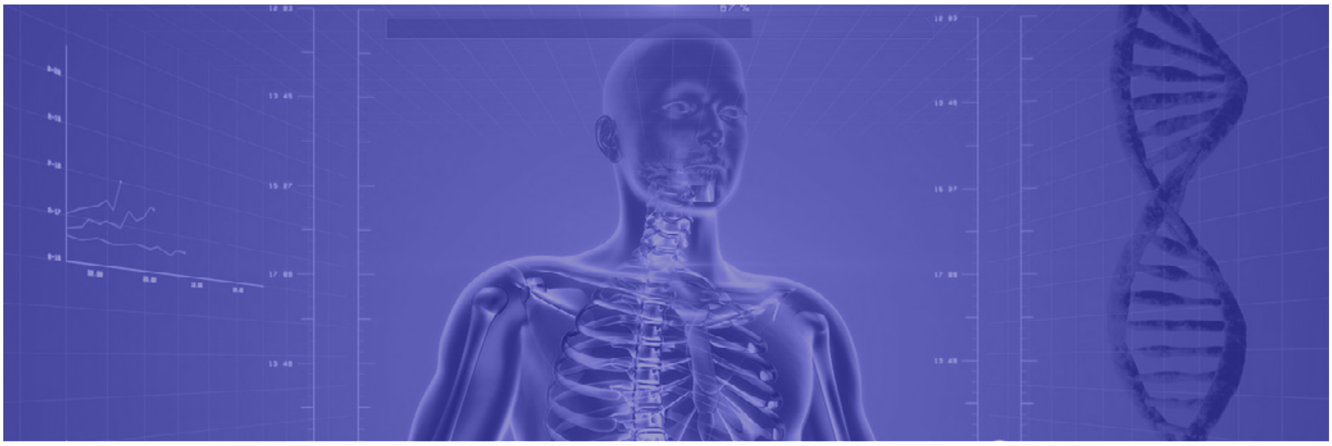
Other Engagement

We continue to use our website as a promotional tool and repository. Over the past year the website was refreshed and restructured to focus on our work, and outcomes such as case studies (genomics-aotearoa.org.nz/our-stories), and news stories and events (genomics-aotearoa.org.nz/news-events). A Vacancies sub-section was added to News & events to promote vacancies in genomics, primarily for the benefit of students and early-career researchers. We also have a Genomics events sub-section for events of interest to our research community. Enhancements were also made to how training information is presented.

Stories to promote the work of our early-career researchers were made public on our Postdoctoral Fellows page (genomics-aotearoa.org.nz/about/postdoctoral-fellowships).

Our researchers also continued their contributions to the Science Media Centre's Sciblogs. As Sciblogs is no longer adding entries, we now have blogs in our website News & events section to fill this gap. Blogging gives our students and early-career researchers an avenue to write for non-technical audiences, promote their work and raise their profiles.

We cross-promote all Genomics Aotearoa activities with our newsletter, which is produced monthly.



Contributing Researchers

Aotearoa NZ Variome - 1803

Stephen Robertson	University of Otago
Phil Wilcox	University of Otago
Lisa Matisoo-Smith	University of Otago
Murray Cox	Massey University
Tony Merriman	University of Otago
Andrew Sporle	University of Auckland
Allamanda Faatoese	University of Otago
Huti Watson	Ngati Porou Hauora

Better Breeding Values - 1805

Dorian Garrick	Massey University
Michael Lee	University of Otago
Phil Wilcox	University of Otago
Shannon Clarke	AgResearch
Rudiger Brauning	AgResearch
David Chagne	Plant & Food Research
Christine Couldrey	LIC
Neville Jopson	AbacusBio
Andrew Hess	AgResearch
Yu Wang	Massey University / LIC
Brittany Jones	University of Otago
Joshua Rich	University of Otago / AgResearch

Bioinformatics - 1811

Peter Dearden	University of Otago
Mik Black	University of Otago
Rudiger Brauning	AgResearch
Kim Handley	University of Auckland
Miles Benton	ESR
Roy Storey	Plant & Food Research
Dan Jones	Plant & Food Research
Carmen Astudillo-Garcia	University of Auckland
Rachael Ashby	AgResearch
Ngoni Faya	University of Otago
Leah Kemp	ESR
Libby Liggins	Massey University
Michael Hoggard	University of Auckland

Bioinformatics - 2104

Mik Black	University of Otago
Rudiger Brauning	AgResearch
Kim Handley	University of Auckland
Miles Benton	ESR
Joep de Ligt	ESR
Rachael Ashby	AgResearch
Ngoni Faya	University of Otago
Leah Kemp	ESR
Libby Liggins	Massey University
Michael Hoggard	University of Auckland
Jian Sheng Boey	University of Auckland

Cell Free Genomics - 1807

Parry Guilford	University of Otago
Mik Black	University of Otago
Fiona Hely	AbacusBio
Jonah Duckles	AbacusBio
Luna Zhang	AbacusBio
Miles Benton	ESR

Clinical Genomics - 1812

Stephen Robertson	University of Otago
Cheng-Yee Chan	CDHB
Natasha Henden	ADHB
Jo Martindale	CCDHB
Padmini Parthasarathy	University of Otago
Kate Gibson	Genetic Health Service NZ
Rachel Stapleton	Genetic Health Service NZ
Ian Hayes	Genetic Health Service NZ
Patrick Yap	Genetic Health Service NZ
Kate Neas	Genetic Health Service NZ
Colina McKeown	Genetic Health Service NZ
Nerine Gregersen	Genetic Health Service NZ

Environmental Metagenomics - 1806

Kim Handley	University of Auckland
Gavin Lear	University of Auckland
Bevan Weir	Manaaki Whenua Landcare Research
Louise Weaver	ESR
Matt Stott	University of Canterbury
Charles Lee	University of Waikato
Susie Wood	Cawthron
David Waite	University of Auckland
Carmen Astudillo-Garcia	University of Auckland
Hwee Sze Tee	University of Auckland
Chanenath Sriaporn	University of Auckland
Jian Shen Boey	University of Auckland
Michael Hoggard	University of Auckland

Environmental Microbiomes - 2101

Kim Handley	University of Auckland
Gavin Lear	University of Auckland
Michael Hoggard	University of Auckland
Jemma Geoghegan	University of Otago
Murray Cox	Massey University
Maui Hudson	University of Waikato
Hwee Sze Tee	University of Auckland
Chanenath Sriaporn	University of Auckland
Jian Shen Boey	University of Auckland
Carmen Astudillo-Garcia	University of Auckland
Annie West	University of Auckland
Esme Spoelstra	University of Auckland
Sophie Palmer Dale	University of Auckland
Millie Ram	University of Auckland

Extending Whole Genome Analysis - 2105

Stephen Robertson	University of Otago
Louise Bicknell	University of Otago
Miles Benton	ESR
Leah Kemp	ESR
Natasha Hendon	ADHB
Jo Martindale	CCDHB
Cheng Yee Chan	CDHB
Kim McGuiness	Rare Diseases NZ
Padmini Pathasarathy	University of Otago
Sam Hawarden	University of Otago

Genome Graphs - 2003

Zoe Yang	ESR
Nuzla Ismail	University of Otago
Mik Black	University of Otago
Joep de Ligt	ESR

Genomic Translational Oncology - 1808

Cris Print	University of Auckland
Richard King	CDHB
Lindsey Harbour	ADHB
Henry Chan	WDHB
Kim Gamet	ADHB
Jaclyn Ting Fowler	ADHB
Chang Ho	WDHB

Genomic Typing of Bacterial Pathogens - 1809

Xiaoyun Ren	ESR
Xochitl Morgan	University of Otago
Christina Straub	ESR
Paul Gardner	University of Otago
Patrick Biggs	Massey University
Joep de Ligt	ESR
Jenny Draper	ESR

High Quality Genomes - 1801

Thomas Buckley	Manaaki Whenua Landcare Research

David Chagne	Plant & Food Research
Jeanne Jacobs	AgResearch
Murray Cox	Massey University
Neil Gemmell	University of Otago
Peter Dearden	University of Otago
Anna Santure	University of Auckland
Emily Telfer	Scion
Maren Wellenreuther	Plant & Food Research / University of Auckland
Chen Wu	Plant & Food Research
Ross Crowhurst	Plant & Food Research
Cecilia Deng	Plant & Food Research
Susan Thomson	Plant & Food Research
Elena Hilario	Plant & Food Research
Richard Newcomb	Plant & Food Research / University of Auckland
Tammy Steeves	University of Canterbury
Gary Houliston	Manaaki Whenua Landcare Research
Duckchul Park	Manaaki Whenua Landcare Research
Patrick Biggs	Massey University
Roy Storey	Plant & Food Research
Orin Mc Cormick	RAPiD Genomics
Matthew Tinning	Australian Genome Research Facility
Alana Alexander	University of Otago
Ann McCartney	Manaaki Whenua Landcare Research
Joseph Guhlin	University of Otago
Roger Moraga	TeaBreak Bioinformatics
Jessie Prebble	Manaaki Whenua Landcare Research
Florian Pichlmueller	University of Otago
Talia Brav-Cubitt	Manaaki Whenua Landcare Research
Manpreet Dhami	Manaaki Whenua Landcare Research
Chris Smith	Manaaki Whenua Landcare Research
Caroline Mitchell	Manaaki Whenua Landcare Research
Emily Koot	Plant & Food Research
Sara Montanari	Plant & Food Research
Seung-Sub (Shane) Choi	University of Auckland

High Quality Genomes and Population Genomics - 2002

Thomas Buckley	Manaaki Whenua Landcare Research
David Chagne	Plant & Food Research
Anna Santure	University of Auckland
Shannon Clarke	AgResearch
Rudiger Brauning	AgResearch
Chris Couldrey	LIC
Jeanne Jacobs	AgResearch
Neil Gemmell	University of Otago
Peter Dearden	University of Otago
Peter Ritchie	Victoria University of Wellington
Maren Wellenreuther	Plant & Food Research
Chen Wu	Plant & Food Research
Amali Thrimawithana	Plant & Food Research

Roy Storey	Plant & Food Research
Susan Thomson	Plant & Food Research
Elena Hilario	Plant & Food Research
Andrew Hess	AgResearch
Charles Hefer	AgResearch
Rachel Ashby	AgResearch
Libby Liggins	Massey University
Michael Lee	University of Otago
Tammy Steeves	University of Canterbury
Annabel Whibley	University of Auckland
Duckchul Park	Manaaki Whenua Landcare Research
Natalie Forsdick	Manaaki Whenua Landcare Research
Alana Alexander	University of Otago
Joseph Guhlin	University of Otago
Tom Oosting	Victoria University of Wellington
Jessie Prebble	Manaaki Whenua Landcare Research
Manpreet Dhami	Manaaki Whenua Landcare Research
Seung-Sub Choi	University of Auckland
Sarah Bailey	University of Auckland
Iggy Carvajal	Plant & Food Research
Emily Koot	Plant & Food Research
Sara Montanari	Plant & Food Research
Cecilia Deng	Plant & Food Research
Rebecca Clarke	AgResearch
Florian Pichlmueller	University of Otago
Lara Urban	University of Otago
Catherine Collins	University of Otago
Anna Clark	University of Otago
Monica Vallender	University of Otago / AgResearch
Vanessa Arranz	Massey University
Imogen Foote	Victoria University of Wellington
Jana Wold	University of Canterbury
Jana Wold	University of Canterbury
Roger Moraga	TeaBreak Bioinformatics

Identifying genetic drivers of *Streptococcus pyogenes* - 2004

Christina Straub	ESR
Paul Gardner	University of Otago
Joep de Ligt	ESR
Julie Bennett	University of Otago
Nigel French	Massey University
Te Whetu Kerekere	University of Otago

Invasomics for Biosecurity - 2102

Angela McGaughran	University of Waikato
Manpreet Dhami	Manaaki Whenua Landcare Research
Ellie Parvizi	University of Waikato
Amy Vaughan	Manaaki Whenua Landcare Research

SING Aotearoa

Maui Hudson	University of Waikato
Phil Wilcox	University of Otago
Katharina Ruckstuhl	University of Otago
Tuti Nikora	University of Waikato
Kimiore Hēnare	University of Auckland

Rakeiora - 1901

Cris Print	University of Auckland
Stephen Robertson	University of Otago
Phil Wilcox	University of Otago
Jennie Harre Hindmarsh	Ngati Porou Hauora
Donia Macartney-Coxson	ESR
Joep de Ligt	ESR
Miles Benton	ESR
Kimiore Hēnare	University of Auckland
Polona Le Quesne Stabej	University of Auckland
Ben Curran	University of Auckland
Martin Kennedy	University of Otago
Hugh Stuart	Ngati Porou Hauora
Frances King	Ngati Porou Hauora
Ben Te Aika	University of Otago
Huti Watson	Ngati Porou Hauora
Ben Halliday	University of Otago
Caroline Koia	Ngati Porou Hauora
Jade Scott	University of Auckland / CTNZ
Jun Huh	NeSI
Rochelle Style	University of Auckland
Helen Wihongi	University of Auckland

Ruatau - 2103

Alana Alexander	University of Otago
Catherine Collins	University of Otago
Simon Hills	Massey University
Kimiore Hēnare	University of Auckland
Karaitiana Taiuru	Independent contractor
Lisa Warbrick	Independent contractor
Phil Wilcox	University of Otago

Te Nohonga Kaitiaki - 1802

Maui Hudson	University of Waikato
Chris Battershill	University of Waikato
Jason Mika	Massey University
Phil Wilcox	University of Otago
Jacinta Ruru	University of Otago
Matt Stott	University of Canterbury
Robert Brookes	University of Waikato
Tuti Nikora	University of Waikato
Ariane Thompson	University of Waikato
Lisa Warbrick	Massey University
Amanda Clarke	University of Waikato
Daniel Patrick	University of Auckland

Glossary of Genomic Terms

TERM	DEFINITION
Assembly	DNA sequencing technology cannot read whole genomes in one go; instead, it reads short pieces of bases from a genomic sequence. Sequence assembly aligns and merges fragments from a longer DNA sequence to reconstruct the original sequence.
Bioinformatics	The science of analysing genomic data.
Candidate genes	Genes of interest related to phenotypes or disease states.
Clinical geneticist	This is a medical doctor with special training in genetics who meets with patients to evaluate, diagnose, and manage genetic disorders. Clinical geneticists also assist in the management of genetic diseases by identifying preventable complications through early and accurate diagnosis and surveillance.
CRISPR-Cas9	CRISPR-Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats) is a method of genetic manipulation consisting of two key molecules that introduce a change into the DNA. The molecules are: An enzyme called Cas9 which can cut strands of DNA at a specific location in the genome so that short sections of DNA can then be added or removed. A piece of RNA sequence called guide RNA (gRNA), located within a longer RNA which guides Cas9 to the correct part of the genome to cut.
DNA methylation	DNA methylation is one type of epigenetic mechanism that modifies the expression of genes and therefore their effects. DNA bases are modified by addition of a methyl group. Methylation can change the activity of a DNA segment without changing the genome sequence.
Epigenetics	Epigenetic mechanisms are those that modify inherited (genomic) gene expression i.e., how, and when those genes are switched on (or off). These epigenetic switches may be due to environmental exposure or lifestyle factors. Non-genetic factors include the environment like diet, gut microbiota, toxin, and drug exposure, psychological and physical stressors, and levels of activity throughout life. Measuring the epigenetic changes that occur in diseases, including cancer and heart disease, can provide understanding of the underlying mechanisms.
Epigenome	The set of all epigenetic modifications to an individual's genome.
Eukaryote	An organism whose cells contain a nucleus surrounded by a membrane and whose DNA is bound together by proteins (histones) into chromosomes. Animals, plants, and fungi are eukaryotes.
Exemplar research	A research model that provides leadership and examples for further research in similar fields.

TERM	DEFINITION
Finishing the sequence	Finishing an assembly involves refining the genomic sequence to eliminate sequencing errors and to close gaps.
Genetic admix/ admixture	The presence of DNA in an individual from a distantly related population or species, because of interbreeding between populations or species who have been reproductively isolated and genetically differentiated. Admixture results in the introduction of new genetic lineages into a population.
Genetic counsellors	These are healthcare professionals with training in human genetics and counselling who guide patients and their whānau with a genetic disorder through the process of understanding and making informed healthcare decisions.
Genetic gain	The rate of genetic improvement within a breeding population over time. An important concept in conventional quantitative genetics and breeding, genetic gain can be defined as the increase in performance achieved annually through artificial selection.
Genetic-linkage mapping	Illustrates the order of genes on a chromosome and the relative distances between those genes.
Genome annotation	The process of attaching biological information to genetic sequences.
Genome duplication	A widespread phenomenon in plant genome evolution, where an organism can sometimes inherit two copies of the genome of its parents, instead of only one copy. The doubling of DNA then persists over generations and the duplicated copies can result in the evolution of new functions.
Genome map	Helps scientists to define which parts of the genome are physically linked to each other. The landmarks on a genome map that aid navigation might include short DNA sequences, regulatory sites that turn genes on and off, and genes themselves.
Genomic signature	Genomic regions of DNA sequences that provide information about the activity of a specific group of genes in a cell or tissue.
Genome variation types	Genome variations include mutations and polymorphisms. Mutation is often used to refer to a variation that is associated with a specific human disease, while the word polymorphism implies a variation that may or may not affect a physical characteristic. Genetic variations also include gene deletions, gene additions and structural variations (see structural variants).
Genome-wide association studies	A search for parts of the genome associated with characteristics of interest, one example being human diseases.
Genotype (noun)	An organism's set of genetic variations.
Genotype (verb)	To determine genetic variation in a genome.
Germline	The sequence of cells which develop into eggs and sperm.
Imputation	The mathematical process of replacing missing data with substituted values.
Introns	Extra sequences of DNA inside genes that are non-coding.

TERM	DEFINITION
Long read sequencing *Also known as:* **Third generation sequencing**	Involves new forms of sequencers that can read long distances down one strand of DNA. There are currently two effective technologies: Pacific Biotechnology (PacBio) - an imaging approach that allows the detection of the incorporation of single labelled base pairs one after another into a strand of DNA being replicated (pacb.com/videos/video-overview-of-smrt-technology) Oxford Nanopore (Nanopore) - uses tiny, charged pores that a strand of DNA is drawn into, and as each base passes through the hole it changes the charge in a way that can be measured (nanoporetech.com/how-it-works). These changes in charge are then assigned to each base and the sequence is built up from there.
Linked read technology	Uses a unique barcode system to label short DNA sequences from individual molecules that are close to each other on the genome, so they can be linked to create longer sequence reads.
Metabolomics	This process detects chemicals or metabolites and provides a read-out of what chemicals are in a tissue at any one time.
Microbiome	The community of microorganisms (such as fungi, bacteria, and viruses) that exists in a particular environment. The environment could be anything from the body of an animal or plant or soils and waterways.
Mutations	The changing of the structure or function of a gene, caused by the alteration of single base units in DNA, or the deletion, insertion, or rearrangement of larger sections of genes or chromosomes. The resulting variant form may be transmitted to subsequent generations.
Non-coding DNA	Genes (coding DNA) account for a small percentage of the DNA in the genome - knowing the entire genome sequence will help scientists study the parts of the genome outside the genes. Non-coding DNA includes the regulatory regions that control how genes are turned on an off, as well as long stretches of DNA of unknown function.
Nonsense mutation	A DNA mutation that results in a non-functional protein.
Nucleotides and bases	A single *base* of DNA is made up of Adenine (A), Cytosine (C), Guanine (G) or Thymine(T). Base pairs are the two opposing nucleotides (a base, a sugar molecule, and a phosphate group) on a double-stranded DNA molecule. Adenine pairs with Thymine, and Cytosine pairs with Guanine.
Omics or **'omics**	Omics is the collective name for the disciplines that characterise and quantify pools of biological molecules that translate into the structure, function, and dynamics of organisms. Examples include genomics, proteomics, metabolomics, metagenomics, phenomics and transcriptomics.
Pangenome	A pangenome (also pan-genome or supragenome) is the entire set of genes from all strains or varieties within a group of organisms.
Phenotype	Observable characteristics influenced both by an organism's genotype and by the environment.
Pathway	In genetics, a pathway is a set of genes that work together in a biological process.
Pharmacogenomics	This is the study of how genomic variation within the individual or their disease (including gene expression, epigenetics, germline, and somatic mutations) influences one person's response to drugs. The aim is to optimise drug therapy by maximising therapeutic effect and minimising adverse effects.
Pipeline	A process for the preparation, development, production, and analysis of genomic data.

TERM	DEFINITION
Population genomics	The large-scale application of genomic information to study populations, including entire genomes of an entire species (e.g., kākāpō).
Prokaryote	Unicellular microbial organism that lacks a nucleus. All bacteria are prokaryotes. A related group are Archaea – these are also unicellular but are different to bacteria.
Proteomics	RNA (made from turned-on genes) is translated into protein. Proteomics is a technique to look at the broad range of proteins in a cell or tissue (using mass spectrometry). We can usually identify which proteins are present in a cell and what they are doing.
RNA	RNA (ribonucleic acid) is the molecule that takes information from DNA to make protein and has many other activities. RNA is only made from genes in the DNA that are turned on.
Short read sequencing *Also known as:* **Next generation** or **second-generation sequencing**	Small segments of DNA strands put in order then assembled in the genome. This is assembled by looking at the sequence of each chunk and finding sequences that overlap (aligning).
SNPs	About 90 percent of human genome variation can be accounted for by single nucleotide polymorphisms, or SNPs (pronounced snips). These are variations that involve just one nucleotide, or base. These markers are used to genotype individuals to better understand population genetic variation and can be used to detect genes that are under selection.
Somatic	Refers to the cells of the body in contrast to the cells that make sperm or eggs.
Structural variation or **Structural Variants (SVs)**	Structural variation describes individual or group differences in genome structure, such as gene deletions, insertions, duplications, inversions, and translocations. These variant regions are scattered throughout genomes and are often associated with gene expression changes and observable differences among individuals (phenotypic differences).
Transcriptomics	Sequencing RNA from a tissue or cell to measure the set of active (expressed) genes.
Variome	The complete set of genetic variations found in populations of species.
Whole genome sequencing (WGS)	The process of determining the complete DNA sequence of an organism's genome.

genomics
aotearoa

Director:
Prof Peter Dearden

Operations Manager:
Jayashree Panjabi

Communications Manager:
Claire Grant

Advisor – Communications and Research:
Jess McLean

Communications Consultant:
Monica Gruber

Training Coordinator:
Dr Tyler McInnes

Vision Mātauranga Manager:
Tracey Godfery

Vision Mātauranga Business Support Coordinator:
Jacqueline Godfery

Training Coordinator (Outgoing):
Ngoni Faya

Genomics Aotearoa would like to thank Monica Gruber, the team at Walsh and Beck, and UniPrint, for their contributions to the production of this annual report.

genomics-aotearoa.org.nz
genomics.aotearoa@otago.ac.nz
youtube.com/c/genomicsaotearoa
twitter.com/genomicsnz

Genomics Aotearoa
Annual Report 2021-2022

ISSN 2816-0088 (Print)
ISSN 2816-0096 (Online)

Printed in New Zealand by UniPrint
Layout and design by Walsh and Beck
Written by Monica Gruber

genomics.aotearoa@otago.ac.nz
genomics-aotearoa.org.nz

This report has been printed on Lenza Green
100% Recycled paper, certified FSC® Recycled

www.genomics-aotearoa.org.nz