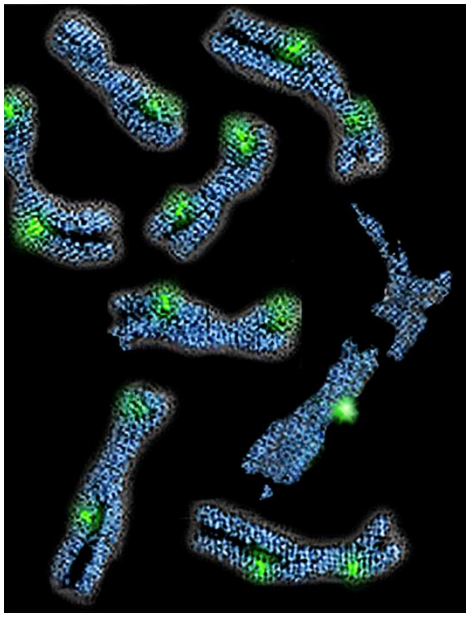
MapNet 2025 Programme




genomics
aotearoa


genomNZ
from agresearch


decode
science


illumina®


BRAGATO
RESEARCH INSTITUTE
RANGAHAU KAREPE, WĀINA O AOTEAROA


康圣序源
Kindstar Sequenon


University
of Otago
ŌTĀKOU WHAKAIHU WAKA


BIOECONOMY
SCIENCE
INSTITUTE
agresearch
āta mātai, mātai whetū


BIOECONOMY
SCIENCE
INSTITUTE
Plant & Food®
Research
Rangahau Ahumāra Kai

Mihimihi

E nga mana, e nga reo, e nga karangataha maha, tēna koutou! Nau mai, haere mai ki te hui nei o MapNet mō te tau 2025! Ko te hui nei i Tauhinu. Ko BSI – Rangahau Ahumāra Kai te whakahaere manaaki. Te ahua o te hui nei kei roto i te whakatauki: **Ka pū te ruha, ka hao te rangatahi.** Ka rongo tātou nga whakaaro mai nga pukenga o nga wa o mua, mai nga wa o muri hoki. Nōreira, e nga kairangahau kua tae mai, nau mai, mihi mai, karanga mai ki a koutou katoa.

Translation: All peoples, all languages, all authorities, greetings! Welcome to MapNet 2025 workshop, held in Tauhinu (Lincoln), hosted by BSI – Plant and Food. The theme for the workshop is embodied in the whakatauki (proverb) '**Ka pū te ruha, ka hao te rangatahi**' which speaks about succession, so we will hear from a some of MapNet's senior researchers, as well as from emerging researchers. So, to all researchers gathered here: greetings and welcome to you all!

About MapNet:

Established in 2005, MapNet is a semi-formal Aotearoa/NZ-wide collective of researchers in various areas of gene mapping and genetical genomics. The key purpose of MapNet is '**encourage pre-commercial, cross-sector, gene mapping related R&D in a range of organisms valuable to NZ's economy, ecology and culture**'. Participation in MapNet is voluntary, at an individual researcher level. All major sectors are typically represented in activities: health, environment, and primary sector (agriculture, horticulture, aquaculture, etc). These activities have included annual workshops that have typically been attended by 50-100 participants, as well as collaborative research activities such as the Virtual Institute of Statistical Genetics, which was funded 2008-2013 and is currently hosted by Genetics Otago. MapNet also holds one-off workshops and has undertaken various advocacy-type activities on behalf of the community of researchers involved in MapNet.

Organising Committee members

Rachael Ashby	Samantha Baldwin	Mik Black	Ting-Hsuan Chen
Rebecca Clarke	Shannon Clarke	Ken Dodds	Jeanne Jacobs
Alastair Lamont	Phillip Wilcox		

Locations

Venue is Fitzgerald room, Plant and Food campus in Lincoln, accessed off the reception area.

Social Functions

Conference dinner is on Tuesday evening from 6:30 at The Laboratory, 17 West Belt, Lincoln.

MapNet Website

MapNet website is located at https://mapnet2025.github.io/

Code of conduct

We are dedicated to providing a harassment-free meeting for everyone, all attendees, speakers, and sponsors are required to abide by the code of conduct located on the MapNet website. Organisers will enforce the code of conduct and expect cooperation to ensure a safe environment for all.

Please respect the requests of speakers and conference attendees that ask or suggest not to be included in social media posts.

Tuesday 11th November	
8:45am	Karakia and Welcome
Keynote	
Session Chair: Samantha Baldwin	
9:00am	**Keynote:** Brian Cullis (University of Wollongong) *DWREML: a computationally efficient R package for fitting aspirational mixed models in genetic improvement programs*
9:45am	**Morning Tea**
Indigenous Health	
Session Chair: Jeanne Jacobs	
10:15am	Phil Wilcox (University of Otago) *Whāia te mātauranga hei oranga mō koutou: Advancing Māori participation and leadership in medical genomics research – an overview*
10:35am	ECR Award: Ben Rangihuna (University of Otago) *Pioneering precision medicine for Māori and Pacific people: Estimating metabolic disease-specific (gout) polygenic risk scores (PRS) to find 'biological hubs'*
10:55am	ECR Award: Clare Adams (University of Otago) *Identifying pleiotropic variants for metabolic conditions in Māori and Pacific peoples*
11:15am	Martin Kennedy (University of Otago, Christchurch) *CYP2D6*71 is a low functioning pharmacogene variant common in Polynesian and Māori people but absent from Europeans*
11:35am	Alastair Lamont (University of Otago) *Population Simulation to Optimise Study Designs and Estimate Polygenic Disease Risk/Resilience in Aotearoa Māori Populations*
11:55pm	**Lunch**
Keynote	
Session Chair: Samantha Baldwin	
1:00pm	**Keynote:** Timothy Bilton (BSI – AgResearch) *A new era of genomic prediction for the primary industries: Genomics for all, phenotyping at scale*
Methods and Teaching	
Session Chair: Alastair Lamont	
1:45pm	Ken Dodds (BSI – AgResearch) *What is the best sequencing depth for genome-wide association studies?*
2:05pm	ECR Award: Julie Blommaert (BSI – Plant and Food) *Can decision trees do the job of GWAS and GBLUP? A case study in Australasian snapper*
2:25pm	John Holmes (University of Canterbury) *Improving breeding value reliability approximation in genetic evaluation with the help of fictitious animals*
2:45pm	Chloé van der Berg (Genomics Aotearoa) *Genomics Aotearoa presents the Bioinformatics Training Programme*
3:05pm	**Afternoon tea**

Animal Disease	
Session Chair: Rebecca Clarke	
3:35pm	John McEwan (BSI – AgResearch) *An Icelandic sheep saga: mapping Bógkreppa to a frameshift mutation in the SHOX gene*
3:55pm	ECR Award: Lucía Mayor Fidalgo (Doñana Biological Station, Seville Spain) *Linkage mapping reveals two candidate regions associated to juvenile idiopathic epilepsy in the Iberian lynx*
4:15pm	Marion Price-Carter (BSI – AgResearch) *ONT SNP amplicon sequencing for rapid and high-resolution Bovine TB strain typing directly from tissue*
Structural Variations	
Session Chair: Ting-Hsuan Chen	
4:35pm	ECR Award: Tram Vi (University of Auckland) *Conserved genomic landscapes of variation in two common invasive birds*
4:55pm	Rebecca Clarke (BSI – AgResearch) *Uncovering the Role of Structural Variants in Genetic Selection for New Zealand Sheep*
5:15pm	Closing remarks for the day
5:30pm	Finish
6:30pm	Dinner – The Laboratory

Wednesday 12th November	
8:25am	Welcome
Keynote	
Session Chair: Samantha Baldwin	
8:30am	**Keynote:** Anna Santure (University of Auckland) *Leveraging two decades of monitoring to map inbreeding depression in the threatened hihi (stitchbird; Notiomystis cincta)"*
Plants Part One	
Session Chair: Phil Wilcox	
9:15am	Annabel Whibley (Bragato Research Institute) *Reading between the vines: genomic and epigenomic variation in Sauvignon blanc*
9:35am	Cen Liau (Bragato Research Institute) *DNA methylation analysis of grapevine with nanopore sequencing*
9:55am	**Morning Tea**
Plants Part Two	
Session Chair: Phil Wilcox	
10:20am	John McCallum (BSI – Plant and Food) *Microhaplotype-based GWAS Reveals Multiple Loci Underlying Diversity of Polyphenol Content in Fruit of Actinidia Hybrids*
10:40am	Margaret Carpenter (BSI – Plant and Food) *Multiplexed amplicon sequencing for DNA fingerprinting in crop breeding programmes*

11:00am	ECR Award: Natalie Graham (BSI – Scion) *Heritability of the tree root microbiome*
11:20am	**MapNet Discussion: Phil Wilcox**
12:00pm	**Lunch**
Aquaculture	
Session Chair: Rachael Ashby	
1:00pm	Jane Symonds (Cawthron Institute) *Climate Adapted Finfish: Salmon Thermotolerance and Multiomics Update*
1:20pm	ECR Award: July Ariñez (University of Otago/ BSI – AgResearch) *Development of genome assembly for climate resilience breeding of farmed Chinook salmon in New Zealand*
1:40pm	Roy Costilla (Cawthron Institute) *Sex prediction using genomic information and machine learning in New Zealand Sockeye salmon*
2:00pm	ECR Award: Mindy Leader (BSI – Plant and Food/ University of Auckland) *Genomics-Informed Breeding for Climate Resilience in Aquaculture*
2:20pm	ECR Award: Megan Scholtens (Cawthron Institute) *Tools to accelerate genetic gain for Ostreid Herpesvirus-1 (OsHV-1) resilience in Pacific oysters (Crassostrea gigas)*
2:40pm	**Closing Remarks and Karakia**
3:00pm	**Finish**

1. Dwreml: a computationally efficient R package for fitting aspirational models in genetic improvement programs.

Cullis, B.R.[1], Smith, A.B.[1], and Butler, D.G.[1]

[1] School of Mathematics and Physics, University of Wollongong, Wollongong, 2522, New South Wales, Australia.

Linear mixed models are widely used in genetic improvement studies. Their primary use has been to facilitate the application of current and appropriate statistical models to phenotypic data-sets, thus assisting with the maintenance of genetic gain. With the advent of genomic technologies the opportunity to maintain genetic gain through the appropriate use of linear mixed models has been significantly challenged with the computational demands of incorporating genomic information in single-step analyses of moderate to large phenotypic panels. This issue is encouraging the use of either implausible statistical models or the use of inefficient multi-step methods in place of the fully efficient single-step approach. This talk presents an overview of DWReml, which is a computationally efficient open source R package for fitting the linear mixed model. DWReml obtains REML estimates of variance parameters using a modified average information algorithm and its efficiency stems from the use of MUMPs, which is a parallel sparse solver, to solve the mixed model equations and obtain key components required in the iterative maximisation of the REML log-likelihood. Its utility for fitting aspirational models is demonstrated using a large multi-environment trial data-set in Chickpeas.

2. Whāia te mātauranga hei oranga mō koutou: Advancing Māori participation and leadership in medical genomics research – an overview

Wilcox, P.L.[1,2]

[1]Department of Mathematics and Statistics, University of Otago, Dunedin, Otago, 9016, New Zealand, [2]Iwi affiliations: Ngāti Rakaipaaka, Rongomaiwaihine, Ngāti Kahungunu ki te Wairoa, Te Aitanga a Mahaki.

Over the last decade our community of researchers has witnessed the advent of Māori-led health genomics research and mātauranga (= Māori knowledge) informed tool development, as well as Māori-centred genomics education initiatives. Key research tools such as the He Kākano Māori Variome and the Rakeiora computational platform have been Māori co-led and designed using tikanga (= Māori ethical frameworks) to guide study protocols. These frameworks have also been successful for partnering with Māori health entities and recruiting study participants, as well as avoiding many of the previous negative experiences of gene technologies and technologists. In parallel, various initiatives have been undertaken to increase te ao Māori content and participation in genomics education, in both Māori and University learning environments. These include the successful Summer Internship of iNdigenous peoples in Genomics Aotearoa (SING-A), which has trained over 100 Māori in genomic technologies, two of whom have gone on to lead nationally important genomics initiatives. In addition, Māori content is being taught in genetics and related papers in at least two Universities in New Zealand, as well as one in the USA. Some of these achievements have been documented by a USA-based Pacific historian. Most recently, we have seen the emergence of Māori-led entities in genomics and related research areas, mirroring overseas exemplars of indigenous-led entities in genomics research. In this presentation I will briefly summarise key developments in this area, existing trends and some future directions, as well as ongoing barriers.

3. Pioneering precision medicine for Māori and Pacific people: Estimating metabolic disease-specific (gout) polygenic risk scores (PRS) to find 'biological hubs'

Rangihuna, B.P.[1], Young, C.K.[1], Lamont, A.[2], Adams, C. I. M.[2], Merriman, T.[3,4], Wilcox, P.[2], Leask, M.P.[1]

[1]Department of Physiology, University of Otago, Dunedin, NZ, [2]Department of Mathematics and Statistics, University of Otago, Dunedin, NZ, [3]Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, US, [4]Department of Microbiology and Immunology, University of Otago, Dunedin, NZ.

Precision medicine, where interventions and treatments are tailored to an individual based on genetic and environmental information, is imminent in health care. Owing to the wealth of information from large biobanks, it promises to improve quality of life, and lower medical costs. However, European-centric genome-wide association studies (GWAS) are known to be less efficacious in predicting polygenic disease risk in other ethnic populations, thus widening the health equity gap for our indigenous Māori and Pacific people. Here, we plan to use statistical methods to identify novel precision medicine targets ('biological hubs') that alter the risk of cardiometabolic disease in Māori and Pacific people. This involves calculating gout-specific PRS using information from a Polynesian-based gout GWAS with Māori and Pacific individuals. Our work has two aims: (1) assess the predictive accuracy of a Polynesian-specific PRS and (2), assess the relationship between the gout PRS and metabolic information (i.e., metabolites) with the goal to find novel pathways that are involved in the pathogenesis of gout. This will form the basis of research focussed on treatment targets. Our results show that the Polynesian gout PRS has predictive capability (AUC (predictive accuracy) = 0.57) which improves when accounting for known gout risk factors (age + sex) and genetic ancestry (AUC = 0.75). Analyses for (2) have not been conducted yet. However, we anticipate that the gout PRS will be associated with metabolites that influence gout onset. Overall, our work has started the process to find novel biological hubs which could aid in treatment development by estimating disease risk for Māori and Pacific peoples. Of equal importance, this research works to address the health inequities experienced by these communities.

4. Identifying pleiotropic variants for metabolic conditions in Māori and Pacific peoples

Adams, C.I.M.[1], Lamont, A.I.[1], Merriman, T.R. [2,3], Cullis, B. [4], Leask, M.[5], and Wilcox, P.L.[1]

[1] Department of Mathematics and Statistics, University of Otago, Dunedin, Otago, 9016, New Zealand., [2] School of Medicine - Immunology and Rheumatology, The University of Alabama at Birmingham, Birmingham, AL, 35294, USA., [3] Department of Microbiology and Immunology, University of Otago, Dunedin, Otago, 9016, New Zealand., [4] School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, New South Wales, 2522, Australia., [5] Physiology, University of Otago, Dunedin, Otago, 9016, New Zealand.

Genomic technologies are foundational to personalized medicine, yet current approaches can often overlook shared genetic architectures and complex co-morbidity. This study employs quantitative genetic approaches to identify pleiotropically acting genes that drive the co-occurrence of phenotypes related to metabolic diseases such as gout, chronic kidney disease, and type 2 diabetes. To do so, we employ the Genetics of Gout, Diabetes, and Kidney Disease (GoGDK) dataset to ensure representation from traditionally underrepresented populations, such as Māori and Pacific Peoples. We explore inter-trait genetic correlations using GBLUP-derived genetic scores to identify trait clusters. From these trait clusters, dimension reduction techniques (e.g. factor analysis) are planned to detect genetic variants affecting linked traits with potential pleiotropic effects via a genome-wide association study (GWAS). This work will provide a statistical framework for understanding metabolic disease trait co-morbidities, potentially contributing to more accurate estimates of genome-wide disease risk for Aotearoa New Zealand populations, and may identify new targets for drug repurposing and discovery. Furthermore, we discuss the future governance and structure of the GoGDK legacy dataset, which implement principles of Indigenous data sovereignty.

5. CYP2D6*71 is a low functioning pharmacogene variant common in Polynesian and Māori people but absent from Europeans

Hitchman, L.M.[1], Magon, N.J.[1], Miller, A.L.[1], Sheen, C.R.[5], Dunn, E.[5], Holloway, B.[1,6], Bozonet, S.M.[1], Pearson, J.F.[2], Faatoese, A.[2], Merriman, T.R.[3,4], Leask, M.[6], Kettle, A.J.[1], Kennedy, M.A.[1]

[1]Department of Pathology and Molecular Medicine, University of Otago, Christchurch, New Zealand; [2]Christchurch Heart Institute, Department of Medicine, University of Otago, Christchurch, New Zealand; [3]Biochemistry Department, University of Otago, Dunedin, New Zealand. [4]Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, Alabama, US; [5] Te Pokapū Auaha Callaghan Innovation, University of Canterbury. Department of Physiology, University of Otago, Dunedin.

CYP2D6 is an important liver protein that metabolises many drugs. The CYP2D6 gene is extremely polymorphic, which can lead to variable activity of CYP2D6, and risks of adverse drug reactions. However, the full extent of variability in CYP2D6 is unknown, particularly for understudied populations. In prior work we showed that an allele called CYP2D6*71 which is not observed in Europeans, constitutes close to 10% of alleles in Māori and Pasifika people. Understanding the functional impact of the CYP2D6*71 allele is crucial to allow accurate inference of drug metabolizer phenotypes.

The key CYP2D6*71 variant (rs118203758) is a G42E substitution in the N-terminal membrane insertion region of CYP2D6. We tested the functional impact of this variant by identifying multiple individuals with CYP2D6*71 homozygote or heterozygous with a known null allele, then using mass spectrometry to detect the metabolic products of solanidine, a recently described biomarker for CYP2D6 activity, in stored plasma samples. In these cases, evidence for limited metabolism of solanidine could be seen, which indicates that CYP2D6*71 is most probably a low-function or poor metabolizer allele. Given the prevalence of this allele in Aotearoa-New Zealand CYP2D6 testing in this country must include CYP2D6*71 to ensure phenotypes are correctly inferred. This work has significant implications for the equitable application of personalised medicine in this country and elsewhere.

6. Population Simulation to Optimise Study Designs and Estimate Polygenic Disease Risk/Resilience in Aotearoa Māori Populations

Lamont, A.[1], Black, M.[2], Wilcox, W.[1]

[1]Department of Mathematics and Statistics, University of Otago, [2]Department of Biochemistry, University of Otago.

For commonly occurring polygenically inherited conditions such as gout, type 2 diabetes, and cardiovascular conditions, disease risk/resilience (DR) estimates have most often been derived from GWAS (genome-wide association studies). Such studies require large sample sizes ($n > 10^4$ participants) genotyped with 10^4-10^7 DNA markers.

These datasets often do not include indigenous peoples, who can have important genetic differences from more commonly represented populations of predominantly European descent. Moreover, existing datasets from Māori (and Pasifika) domiciled in New Zealand are few, and those that could be utilised consist of fewer than two thousand individuals - and thus are underpowered for clinically accurate disease risk/resilience prediction. In addition, establishing sufficiently large GWAS is unlikely in Aotearoa/NZ because of substantive costs associated with generating genotypic data and reluctance of many Māori to participate in such studies.

In order to offset further health inequities arising from a lack of Māori-specific DR prediction models, new studies are required. Such studies require both (a) optimal designs that incorporate known genetic relationships on non-genotyped as well as genotyped individuals, and (b) analytical methods that more accurately predict phenotype than GWAS-based methods such as polygenic risk scores. Well-established quantitative genetics approaches such as ssGBLUP are not commonly used in DR prediction, but are well-suited for this purpose with appropriate datasets.

We have used a population simulator (SLiM) to model genetic structures of Māori communities (i.e., whānau/hapū/iwi), incorporating estimates of effective population sizes prior to European admixture, as well as post-colonisation admixture with Europeans. We are using these simulations to explore what features of study design and analytical methods lead to optimal disease risk/resilience prediction.

I will describe the simulations and compare how choices of study design / analytical method affect prediction performance, plus implications for potential real-world studies.

7. A new era of genomic prediction for the primary industries: Genomics for all, phenotyping at scale

Bilton, T.P.[1], Dodds, K.G.[1], Clarke, S.M.[1], Johnson, P.L.[1], Schofield, M.R.[2], Black, M.A.[3], Jacobs, J.M.E[4], Perry, B.J.[1], van Stijn, T.C.[1], Henry, H.[1], McRae, K.[1], Hickey, S.M.[5], Jonker, A.[6], Griffiths, A.G.[6], McEwan, J.C.[1], Rowe, S.J.[1]

[1]Bioeconomy Science Institute, Invermay Agricultural Centre, Mosgiel, NZ, [2]University of Otago, Department of Mathematics and Statistics, Dunedin, NZ, [3]University of Otago, Department of Biochemistry, Dunedin, NZ, [4]Bioeconomy Science Institute, Lincoln Research Centre, Lincoln, NZ, [5]Bioeconomy Science Institute, Ruakura Research Centre, Hamilton, NZ, [6]Bioeconomy Science Institute, Grasslands Research Centre, Palmerston North, NZ.

Genomic prediction utilizes genomic information to predict traits in individuals and can provide greater accuracy than using pedigree information alone, facilitating faster genetic progress in breeding programs. New technologies developed over the past decade have enabled low-cost, high-throughput methods to be developed for both phenotyping and genotyping (e.g., reduced representational sequencing) making genomic prediction, more accessible to low-resourced species. One component that is essential for genomic prediction is having some way to describe the genetic relationships between individuals. In diploid individuals, methods for determining genetic relationships are well developed but for species with complex inheritance patterns (i.e., polyploids) or breeding programs where samples are pooled during sequencing for efficiency (i.e., phenotyping at the pool level), appropriate tools and theory are lacking. Here, we present methods that appropriately estimate genetic relationships for polyploid species and pooled samples that also account for the complexity of data generated using high-throughput sequencing methods. Another component that is important for genomic prediction is having sufficient phenotypic data for individuals in the reference set to train the prediction models. Some traits, however, can be very difficult and expensive to measure directly and finding "proxy" phenotypes that are heritable and genetically correlated with the trait of interest provides an alternative phenotyping strategy. Traits where researchers are searching for novel proxy measures are methane emissions and feed efficiency in livestock. To date, fatty acid and electromagnetic spectrum profiles from milk and meat samples, and rumen and oral microbiome profiles have shown great promise as proxy measures of methane emissions and feed efficiency in sheep and cattle. We discuss current research in this area and potential future directions.

8. What is the best sequencing depth for genome-wide association studies?

Dodds, K.G.[1], McEwan, J.C.[1], Bilton, T.P.[1], Brauning, R.[1], Clarke, S.M.[1]

[1]Bioeconomy Science Institute, Invermay Agricultural Centre, Mosgiel, NZ.

When planning a genome-wide association study (GWAS) using sequencing-based marker platforms, researchers need to decide how much resource to spend on sequencing each individual versus including more individuals in the study. If sequencing effort per individual is high, there will be high confidence in the genotypes, but this allows few individuals. Conversely, with many individuals the same budget would require low-depth sequencing, resulting in genotype uncertainty. We considered a situation where the sequencing effort, defined as the product of the mean number of reads for a genotype and the number of individuals genotyped, is held constant. A set of unrelated individuals segregating for a quantitative trait locus (QTL) was simulated after which datasets where the alleles were sampled from each QTL genotype with differing depths were formed. Genotype probabilities and allelic dosage (expected number of reference alleles) were calculated using priors with dataset allele frequencies and Hardy-Weinberg equilibrium. Individual phenotypes were simulated by adding a residual to their QTL genotype value. Association tests, regressing phenotypes on allelic dosage (additive model) or on the genotype probabilities (genotypic model) were applied to give the power of each study design. We found that low sequencing effort, even less than one read per genotype on average, gave the highest power, except when there was heterozygote advantage (overdominance) in which case the optimal depth was around 2-6. These results will help inform study design. In many situations this will result in a design with many individuals, and which forgoes obtaining accurate genotype calls.

9. Can decision trees do the job of GWAS and GBLUP? A case study in Australasian snapper

Blommaert, J.[1], Vander Velpen, S.[1,2], Bayer, P.[3], Catanach, A.[1], Jesson, L.[1], Wellenreuther, M.[1,4]

[1] Bioeconomy Science Institute, Nelson Research Centre, Nelson, New Zealand, [2] Hogeschool Leiden, Leiden, The Netherlands, [3] Minderoo Foundation, Perth, Australia, [4] The School of Biological Sciences, The University of Auckland, New Zealand.

The corner stones of molecular breeding, genome wide association studies (GWAS) and genomic best linear unbiased prediction (GBLUP), each tackle different but complimentary aspects of the breeding process. GWAS aims to identify variants associated with certain traits, whereas GBLUP aims to predict phenotypes of breeding individuals. Both are complimentary, but here, we tested whether decision trees can both predict phenotypes and identify important variants using genomic data from the fourth generation of Australasian snapper (*Chrysophrys auratus*) in the long-running breeding programme at BSI Nelson.

We applied high-throughput, image-based phenotyping to quantify 13 morphological traits alongside manual weight and fork length. Heritabilities were estimated, and GWAS revealed 24 SNPs associated with growth, many near genes involved in metabolism and development. XGBoost models trained on SNP genotypes achieved moderate predictive power, with substantial overlap between SNPs prioritised by GWAS and those highlighted by machine learning. Thus, decision trees captured similar signals while providing direct prediction of phenotypes.

We further extended the comparison to structural variants (SVs), which are often underrepresented in GWAS and GBLUP. To address this, we incorporated SVs into GWAS and ML approaches used with SNPs and also applied a custom SV-focused method to identify growth-associated variants. This custom method identified more growth-associated variants than either GWAS or ML, suggesting that inclusion of SVs can substantially enhance trait discovery when biologically meaningful analyses are applied even though other methods to study them are less mature than those for SNPs.

Overall, decision trees recovered overlapping signals to those identified by GWAS, and provided some predictive utility, and highlighted complementary sets of variants. These results suggest that tree-based machine learning is a valuable addition to the aquaculture breeding toolbox, but other data science tools to reduce dataset dimensionality may help improve its application to genomic prediction.

10. Improving breeding value reliability approximation in genetic evaluation with the help of fictitious animals.

Holmes, J.B.[1], Lee, M.A.[2,3]

[1]Department of Mathematics and Statistics, University of Canterbury, Christchurch, NZ, [2]Department of Mathematics and Statistics, University of Otago, Dunedin, NZ, [3] Beef and Lamb Genetics, Dunedin, NZ.

Breeding values are typically estimated by fitting a linear mixed model, $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{e}$; $\boldsymbol{u} \sim N(\boldsymbol{0}, \sigma_g^2\boldsymbol{G}), \boldsymbol{e} \sim N(\boldsymbol{0}, \sigma_e^2\boldsymbol{R})$, to trait data. In order to find the breeding value reliability for animal i, $r_i^2 = \frac{Var(\hat{\boldsymbol{u}}_i)}{Var(\boldsymbol{u}_i)} = 1 - \frac{Var(\hat{\boldsymbol{u}}_i - \boldsymbol{u}_i)}{Var(\boldsymbol{u}_i)}$, we need to invert a $q \times q$ matrix, where q is the number of random effect levels. In animal breeding contexts, where, in a single trait model, q would be equal to the number of animals with pedigree and/or genomic data, direct inversion is too costly, meaning reliabilities are approximated. While fast and accurate methods to approximate reliability have been developed in contexts where relatedness, $\boldsymbol{G}$, is estimated using pedigree data alone, these approaches break down when $\boldsymbol{G}$ is estimated using (partial) genomic data.

In the context where $\boldsymbol{G}$ is estimated using genomic data, Ben Zaabza, Mäntysaari, and Strandén proposed approximating reliabilities based on drawing random samples, $\boldsymbol{a}_j \sim N(\boldsymbol{0}, \boldsymbol{G}); j = 1, 2, \dots, d$, then fitting the model $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{A}\boldsymbol{g} + \boldsymbol{e}$, where $\boldsymbol{A} = (\boldsymbol{a}_1\ \boldsymbol{a}_2\ \cdots \boldsymbol{a}_d)$ with $d << q$ random effect levels and $Var(\hat{\boldsymbol{u}}_i - \boldsymbol{u}_i)$ approximated with $\boldsymbol{A}_i Var(\hat{\boldsymbol{g}} - \boldsymbol{g})\boldsymbol{A}_i'$. While this resulted in approximate reliabilities that were highly correlated with the true reliabilities, the approximate reliabilities were biased upwards, especially when the true reliability was low, and/or d was small.

We will demonstrate if we augment the model with fictitious animals, such that $\begin{pmatrix}\boldsymbol{y}\\ \boldsymbol{y_f}\end{pmatrix} = \begin{pmatrix}\boldsymbol{X}\\ \boldsymbol{0}\end{pmatrix}\boldsymbol{\beta} + \begin{pmatrix}\boldsymbol{Z} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{D}\end{pmatrix}\begin{pmatrix}\boldsymbol{u}\\ \boldsymbol{u_f}\end{pmatrix} + \begin{pmatrix}\boldsymbol{e}\\ \boldsymbol{e_f}\end{pmatrix}$, with $\boldsymbol{u}$, the true animals, and $\boldsymbol{u_f}$, the fictitious animals, independent and $\boldsymbol{D}$ a diagonal matrix, we will know the true reliabilities of $\boldsymbol{u_f}$ analytically and thus be able to de-bias the approximate reliabilities for $\boldsymbol{u}$ at minimal additional cost. We will then use the fact the approximation of $\boldsymbol{G}$ is Wishart distributed to estimate standard errors of the approximate reliabilities, thus determining a minimum value for d. Put together, our modifications suggest it is possible to obtain an unbiased estimate of breeding value reliability with low uncertainty at around $d = 10{,}000$.

11. Genomics Aotearoa presents the Bioinformatics Training Programme

van der Burg, C.A.[1,2], Dearden, P.K. [1,2] and Black, M.A. [1,2]

[1]Genomics Aotearoa, [2]Department of Biochemistry, University of Otago, Dunedin, NZ.

Genomics Aotearoa is a collaborative research platform which brings researchers together to build infrastructure and develop national capability in genomics and bioinformatics. The Bioinformatics Training Programme (BTP), run in partnership between Genomics Aotearoa and NeSI from 2019-2025, and with REANNZ from mid-2025, aims to upskill the research community of Aotearoa in the rapidly growing field of bioinformatics, and also in the use of high-performance computing to accelerate their research. Here we present an overview of our Training Programme, including the range of workshops we offer from beginner through to advanced level bioinformatic training, and future plans to continue to grow this initiative. Since its inception in 2019, the Training Programme has hosted more than 120 workshops and reached more than 2,000 participants. Looking ahead, we aim to expand the Programme, developing new workshops to meet evolving research needs, broadening our teaching team and strengthening connections within the Aotearoa research community. With an emphasis on inclusivity and accessibility modeled on The Carpentries international training organization, our long-term goal is to ensure that New Zealand researchers are equipped with the skills, confidence, and community needed to thrive in genomics and bioinformatics, ultimately driving innovation and discovery across the country.

12. An Icelandic sheep saga: mapping Bógkreppa to a frameshift mutation in the SHOX gene

McEwan, J.C.[1], Einarsson, E.[2], Sigurðsson, Á.[3], Clarke, R.[1], Oddsdóttir, C.[4], Baird, H.[1], Brauning, R.[1], Clarke, S.[1], Sveinsson, S.[5]

[1]AgResearch, Invermay, Mosgiel, NZ. [2]RML, Höfðabraut, Hvammstangi, Iceland. [3]Decode Genetics, Reykjavik, Iceland. [4]University of Iceland, Keldur, Iceland. [5]Matís.Vínlandsleið, Reykjavík, Iceland.

The Icelandic sheep breed has been isolated for over 1,000 years and is renowned for its adaptability to harsh climates and its unique dual-coated fleece. Rapid genetic progress has been achieved over the last 30 years through a combination of BLUP-based breeding, particularly for carcass meat yield, and the extensive use of artificial insemination. However, in recent years, lambs in rare cases were being born with Bógkreppa, or "deformed and crooked legs", and these were descended from a few high-index AI rams that had been used extensively. Runs of homozygosity mapping with the ISGC 606K marker HD sheep chip narrowed a putative recessive mutation to a 753Kbp region on the pseudo-autosomal region (PAR) of the X chromosome, which contained 4 genes, including SHOX. Mutations affecting this gene in humans included, Turner Syndrome, Madelung deformity and Leri-Weill dyschondrosteosis. Phenotypic manifestations in humans were generally more severe in females, and in males, muscular hypertrophy was a frequent finding. A unique mutation observed only in Icelandic sheep, and in high linkage disequilibrium to the mutation, was detected adjacent to a SNP on the HD chip that could also be scored using the HD chip. This mutation was used in the interim to classify putative carriers while the putative homozygotes were sequenced using Illumina and ONT technology. These confirmed results obtained independently using ABI sequencing, which indicate that the causative mutation is almost certainly a deletion causing a frameshift mutation in the SHOX gene. Concurrently, the Icelandic sheep breeders are screening individuals for rare variants that confer scrapie resistance and will combine these variants with others in a genomic selection programme using the GenomNZ 80K chip. The putative origin of the mutation and its effect on production traits in the heterozygous state are still under investigation.

13. Linkage mapping reveals two candidate regions associated to juvenile idiopathic epilepsy in the Iberian lynx

Mayor-Fidalgo, L.[1], Kleinman-Ruiz, D.[1], Bazzicalupo, E.[1], Godoy, J.A.[1]

[1]Departament of Ecology and Evolution, Doñana Biological Station, Seville, 41092, Spain.

The Iberian lynx (*Lynx pardinus*) is an endangered felid endemic to the Iberian peninsula that has undergone various severe bottlenecks, leaving multiple signs of genetic erosion in the species, including reduced genome-wide diversity, elevated inbreeding coefficients and accumulation of genetic load. Among the conservation actions implemented, a captive breeding program was established, which has been successfully maintaining the remaining genetic variation. Nonetheless, since the start of the program, 20 cubs have been diagnosed with juvenile idiopathic epilepsy, a condition with a presumed genetic origin. Epilepsy has been estimated to have high heritability, a recessive mode of inheritance, and to be potentially driven by one or few major-effect loci. The mating between suspected carriers has been avoided since 2017, resulting in no new cases at the cost of restricted genetic management options.

The primary goal of this work is to elucidate the genetic architecture underlying epilepsy in the species. We applied a low-coverage whole-genome sequencing (lcWGS) approach combined with GLIMPSE1 imputation to generate genome-wide SNP data from affected individuals and their relatives to conduct linkage analyses. Specifically, we implemented a parametric recessive model of inheritance on nuclear families with the MERLIN software, complemented using *loki*, a Bayesian framework that enables the integration of complex pedigrees.

Preliminary results revealed two regions located on chromosomes A2 and B2 showing evidence of linkage. Ongoing work aims to refine these candidate regions and integrate complementary methodologies, such as GWAS or homozygosity mapping. Identifying the putative causal variant(s) or tightly linked markers would allow their integration into the genetic management of the species. In practice, this could enable the detection of true carriers in the captive breeding population allowing more informed genetic management decisions that avoid unnecessary restrictions on breeding and help preserve genetic diversity.

14. ONT SNP amplicon sequencing for rapid and high-resolution Bovine TB strain typing directly from tissue

Price-Carter, M.[1], Sage, R.[1]

[1]BSI, AgResearch, Hopkirk Research Institute, Manawatu, NZ.

Problem. Bovine tuberculosis (bTB) is a zoonotic disease caused by *Mycobacterium bovis*. Because it is maintained in both New Zealand (NZ) livestock and wildlife it has been challenging to control. NZ *M. bovis* strain types are correlated with geographical location and strain typing is used by OSPRI (TB free) to guide how to best manage new infections. Current *M. bovis* genotyping requires culture from crude tissue homogenate in a PC3 facility. Extracted culture DNA is then outsourced to an Illumina sequencing facility for whole genome sequencing (WGS). This process can take up to 12 weeks which can impact the ability of OSPRI to track transmission during outbreaks thus impacting NZ farmers. Our aim is to develop in-house Oxford Nanopore Technology (ONT) genotyping of *M. bovis* directly from infected tissue.

Brief Methods. Assays were designed based on PCR amplified genomic fragments (amplicons) containing single nucleotide polymorphisms (SNPs) that are shared by strains from the four major NZ clades, important minor sub-clusters, and recent outbreaks. Eight sets of compatible primers were designed for PCR amplification in multiplexed groups of three to five which were pooled and sequenced on ONT Flongle flow cells. Trimmed sequencing reads were mapped to an *M. bovis* reference genome to detect the targeted SNPs.

Conclusions. Because of the paucibacillary nature of most *M. bovis* infections, we have so far been unable to genotype by direct WGS from tissue but can now rapidly genotype high and low *M. bovis* load tissue samples by ONT sequencing SNP amplicons based on high resolution WGS data. With the current SNPs set, this assay, from start to reporting results takes six business days to complete. Side by side comparisons on incoming requests will help to define its utility relative to WGS.

15. Conserved genomic landscapes of variation in two common invasive birds

Vi, T.[1], Stuart, K.[1,2], Whibley, A.[1,3], Lloret-Villas, A.[4], Santure, A.W.[1]

[1]School of Biological Science, University of Auckland, Auckland, New Zealand, [2]Applied BioSciences, Macquarie University, Sydney, Australia, [3]Bragato Research Institute, Lincoln, New Zealand, [4]Globe Institute, University of Copenhagen, Copenhagen, Denmark.

There is increasing evidence that structural variants (SVs) and transposable elements (TEs) play important roles in adaptation to novel environments. Avian genomes are characterized by a highly conserved genome structure and a significantly lower proportion of repetitive elements. However, the importance of their unique genomic landscapes to the evolutionary potential of species has not been fully understood, particularly for invasive bird species characterized by their ability to adapt and spread rapidly in new environments. In this study, we explore and compare the SV and TE profiles of two invasive birds of the Sturnidae family in New Zealand, common myna (*Acridotheres tristis*) and common starling (*Sturnus vulgaris*). Their similar genomic structure, with a high portion of syntenic regions, but different demographic histories, make these two species a valuable case study for studying similarities and differences in their SV and TE landscapes across the genome. We used Oxford Nanopore Technologies (ONT), which allows detection of SVs and TEs along with SNPs, to sequence 16 and 15 samples of myna and starling, respectively. These data offer an opportunity to assess the impact of SVs in driving genetic diversity of each species using population genetic analyses, and to analyze the role of TEs in genomic structure by integrating variation and methylation landscapes. Our preliminary results in SV profiling show that SV abundance is similar in the syntenic blocks of both species, however, is higher in non-syntenic regions. Genetic diversity driven by SVs is lower in myna compared to starling, as expected from stronger bottlenecks in myna. These and further results of this work will help us gain a deeper insight into the invasive genomics of birds, crucial for their invasion management.

16. Uncovering the Role of Structural Variants in Genetic Selection for New Zealand Sheep

Clarke, R.M.[1], Brauning, R.[1], Baird, H.[1], Dodds, K.G.[1], Lee, M.A.[2,3] & Clarke, S.M.[1]

[1]AgResearch, Invermay, Mosgiel, NZ, [2]Beef + Lamb New Zealand Genetics, Dunedin, NZ, [3]Department of Mathematics and Statistics, University of Otago, Dunedin, New Zealand.

Sheep farming is essential to New Zealand's economy, with 23 million sheep producing over 4.3 billion $NZD in export revenue in 2024. This industry faces economic, environmental and social challenges with genetic improvement having the potential to help address these. Currently single-nucleotide polymorphisms (SNPs) are the main contribution to genomically selecting animals however structural variations (SV) play a major role in genetic diversity and phenotypic variations, yet they remain largely unexplored in most domesticated animals. Compared to SNPs, SV can have a larger effect on phenotype and gene expression. The aim of this research is to identify SV that impact traits of importance and that could be incorporated into breeding values for improved genetic gain.

SNP arrays can be used to determine SV by utilizing the intensity values derived from each sample. Commercial SNP arrays are low cost and high density providing an excellent resource for SV detection. The International Sheep Genomics Consortium Ovine High-density chip (HD-SNP) contains 606,006 SNPs, with thousands of animals (28,569) genotyped on this chip. This population also contains phenotypic data for meat yield and quality traits.

EnsembleCNV was used to detect copy number variations regions (CNVR) with 15,292 identified. 7,414 uniquely identified CNVR were found and all CNVR were analysed for inheritance, possible phenotypic effects based on genes impacted and overlap with known QTL regions. Adjusted phenotypes were regressed on contemporary groups and fixed effects assuming a model and data obtained from Beef+Lamb NZ Genetics. The residuals after fitting these models were used in a linear regression with the CNVR. Significant CNVs were determined with a p-value < 0.05 after FDR correction. From the 24 traits investigated 1,138 CNVR were found to be significantly associated with a trait. This ranged from 0-603 CNVR per trait. This is one of the first studies to investigate SV in a large-scale New Zealand sheep population in association with phenotypic traits that are important for breeding values.

17. Leveraging two decades of monitoring to map adaptive potential and inbreeding depression in the threatened hihi (stitchbird; *Notiomystis cincta*)

Santure, A.W.[1], Tan, H.Z.[1], Duntsch, L.[1], Bailey, S.[1], Guhlin, J.[2], Lee, K.D.[1], Rutschmann, A.[1], Scherer, P.[1], Senanayake, D.S.[3], Vi, T.[1], Brekke, P.[4], de Villemereuil, P.[1], Ewen, J.G.[4], Stuart, K.C.[1], Whibley, A.[1]

[1]School of Biological Sciences, Waipapa Taumata Rau - University of Auckland, Auckland, NZ, [2]Department of Biochemistry, Ōtākou Whakaihu Waka - University of Otago, Dunedin, NZ, [3]New Zealand eScience Infrastructure(NeSI), Waipapa Taumata Rau - University of Auckland, Auckland, New Zealand, [4]Institute of Zoology, Zoological Society of London, UK.

Population declines and fragmentation have resulted in an increasing number of small and isolated populations. Small populations face the loss of genetic diversity and increased inbreeding. While it is assumed that these will impact individual survival and reproduction (i.e., fitness) and reduce the adaptive potential of the population, few studies have been able to measure fitness directly and hence link genomic diversity to fitness in the wild. Long-term, individual-based studies provide powerful case studies to better understand evolutionary processes and responses, and to assess whether genomic measures can act as a proxy in the absence of fitness data. Hihi (stitchbird, *Notiomystis cincta*) is a threatened endemic Aotearoa New Zealand passerine. Using extensive lifetime fitness data from the Tiritiri Matangi island population dating back to 2005, and combined with genome sequencing information for 431 birds, we have used a number of complementary approaches to measure hihi adaptive potential. These have included genetic diversity measures and directly inferring the heritability of fitness, and also approaches that assess inbreeding depression in the species via the inbreeding-fitness relationship with both structural variants and SNPs, a genome wide association for fitness, and using genomic information to infer harmful variants in the genome. I'll discuss what our results mean in terms of how best to preserve this taonga species, and how we might use genomic information to understand the evolutionary potential of other threatened species.

18. Reading between the vines: genomic and epigenomic variation in Sauvignon blanc

Whibley, A.[1,2], Liau, Y.[1], Hill, A. M.[1], Lizamore, D.[1]

[1]Grapevine Improvement, Bragato Research Institute, Lincoln, NZ, [2]School of Biological Sciences, University of Auckland, Auckland, NZ.

A long history of domestication, clonal selection, and vegetative propagation has left its mark on the grapevine genome. Recent advances in sequencing technologies and bioinformatics now enable us to interrogate this genome biology with unprecedented resolution. Understanding the functional significance of genetic and epigenetic variation is a critical foundation for identifying molecular targets and developing genomics-informed strategies to support precision breeding. We are putting this into practice with Sauvignon blanc, the flagship grapevine variety of the New Zealand wine industry, ultimately aiming to expand the grower toolkit by providing new clonal material with enhanced performance and resilience under future growing conditions.

I will outline our approaches to characterising genomic and epigenomic variation among existing and newly generated Sauvignon blanc clones. Leveraging long-read sequencing datasets from Oxford Nanopore Technologies and a high-quality diploid reference assembly, we are cataloguing variation and gaining insights into structural complexity of nuclear and organelle genomes, somatic mosaicism, and mutational dynamics. I will also highlight methodological challenges, many of which come from the application of tools that were optimised in genomic systems with quite different properties, and present some of the analytical solutions we have implemented.

19. DNA methylation analysis of grapevine with nanopore sequencing

Liau, Y.[1], Whibley, A.[1,2], Vanga, B.R.[1], Hill, A.[1], Lizamore D.[1]

[1]Grapevine Improvement, Bragato Research Institute, Lincoln, NZ, [2]School of Biological Sciences, University of Auckland, Auckland, NZ.

Epigenetic regulation, including DNA methylation, plays an important role in plant development and environmental response. Recent advancements in genomic technologies have enabled more comprehensive and direct analysis of DNA methylation patterns. The Grapevine Improvement team at the Bragato Research Institute is using Oxford Nanopore sequencing to evaluate DNA methylation changes in grapevine associated with environmental variability, stress treatments, and somaclonal variations.

DNA methylation signals in cytosine and adenosine residues were obtained alongside canonical base information using the ONT Dorado basecaller. DNA methylation frequency at each site was extracted using ModKit. We compared the technical performance of different Dorado basecalling versions and performed differential methylation analysis across our samples.

There was a strong correlation between different dorado versions for DNA methylation in the CG ($R = 0.98$-0.99) and CHG ($R = 0.95$-0.97) contexts. However, the correlation for CHH methylation was more variable, ranging from $R = 0.41$ to 0.79. When comparing against whole-genome bisulfite sequencing, we observed good agreement for CG and CHG methylation ($R > 0.9$) and lower agreement for CHH ($R = 0.2$-0.5). Analysis of methylation patterns across different environmental conditions revealed notable variation in the epigenetic landscape with distinct methylation profiles observed in grapevine plants grown under different environmental conditions. Somaclones derived from stress-induced embryogenic callus exhibited significant epigenetic changes, which tended to stabilize over time as the plants matured. During early developmental stages, such as in embryogenic callus and juvenile plants, higher methylation levels were observed in the CHH context, with methylation levels decreasing as the plants transitioned into more mature states.

20. Microhaplotype-based GWAS Reveals Multiple Loci Underlying Diversity of Polyphenol Content in Fruit of *Actinidia* Hybrids

McCallum, J.A. [1], Joyce, N. [1] Millar, T. [1], Robson T [1], Heywood, A. [1], Jesson, L.[1], Catanach, A. [1], Guo L.[1], Macknight, R.[2], Patrick, E. [2], Datson, P.[2], Currie, A.[2]

[1]Bioeconomy Research Institute, NZ ,[2] Kiwifruit Breeding Centre, NZ.

Kiwiberry (*Actinidia arguta* var. *arguta)* is a small-fruited relative of kiwifruit (*A. chinensis*) that is rich in polyphenols and other healthful phytochemicals but is early in domestication. To understand the genetic architecture of polyphenol composition in kiwiberry relating to health and taste we conducted semi-targetted LCMS profiling of skin tissue extracts from a cohort of seven connected autotetraploid families originating from hybrids between *Actinidia arguta* var. *arguta, A. arguta* var. *purpurea* and *A. melanandra*. These families were genotyped with a 3k FlexSeq array and GWAS was conducted using methods based on SNP and microhaplotype markers. GWAS revealed multiple QTL with strong and distinct effects on specific polyphenol classes. Haplotype-based methods enabled inference of haplotype origins and development of kmer-based tags for wider analyses and prediction. These findings provide a basis for understanding the extensive chemodiversity observed in kiwiberry and fixing desirable allelic combinations in breeding for human health and taste attributes.

21. Multiplexed amplicon sequencing for DNA fingerprinting in crop breeding programmes

Carpenter, M.A.[1], Chagné, D.[2], Wiedow, C.[2], Johnston, P.A[1], Bloomer, R.[1], Grande, A.[2], Drury, H.[2], Pither-Joyce, M.[1], Bowatte, D.[2], Wright, K.[1], Cao, Z.[1], Meiyalaghan, V.[1], Catanach, A.[1], Thomson, S.[1], Deng, C.[3], Barrell, P.[1], Montanari, S.[4], Monaghan, K.[1], David, C.[1]

New Zealand Institute for Bioeconomy Science, [1]Lincoln, [2]Palmerston North, [3]Auckland, [4]Motueka, New Zealand.

Fingerprinting using DNA markers plays a crucial role in plant breeding, enabling true-to-type testing, pedigree verification, and marker-assisted selection (MAS). Traditionally, this has been accomplished using simple sequence repeats (SSRs), but this technology is gradually being phased out in favour of Next Generation Sequencing (NGS). NGS offers more detailed sequence information than SSRs, and advances in sequencing technology have made it increasingly cost-effective. The MinION™ portable nanopore sequencing device is well-suited for affordable sequencing on a smaller scale, providing rapid turnaround times that align with the needs of breeding programmes.

We are currently developing methods to transition from SSR analysis on fragment analysers to amplicon sequencing using the MinION. This effort includes testing the feasibility of sequencing existing SSR amplicons on the MinION, as well as designing new sets of multiplexed amplicons that target single nucleotide polymorphisms (SNPs). Additionally, we are creating user-friendly bioinformatics pipelines for automated sequence data processing, which require minimal bioinformatics expertise and can be easily operated by laboratory staff.

To date, we have successfully trialled MinION sequencing across five crops as a proof of concept. This included a set of 96 multiplexed amplicons for potato, a set of 170 amplicons for oats, and sequencing SSR amplicons for apple, hops and kiwifruit. We are now expanding these fingerprinting systems to include an additional nine crops. For some crops the amplicon panels will incorporate markers associated with specific traits to enable MAS in our breeding programmes.

22. Heritability of the tree root microbiome

Graham, N.J.[1,2], Day N. J.[2], Slavov, G.[3, 4], Klápště. J.[5], Wakelin, S.A.[1]

[1]Scion Group, Bioeconomy Science Institute, Rotorua, NZ, [2]School of Biological Sciences, Victoria University of Wellington, Wellington, NZ, [3]Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, UK, [4]Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR, USA, [5]Tree Breeding Australia Ltd., Mount Gambier, Australia.

Long-lived trees such as conifers may rely on microbial partners to adapt to changing environments, yet host–microbe interactions remain poorly understood in most species. We investigated the role of host genetics in shaping root-associated microbiomes of *Pinus radiata* D. Don, a commercially important conifer. Root samples from 528 individuals in a clonal breeding trial were profiled using bacterial and fungal amplicon sequencing, and microbiome structure was analysed in relation to host genetic factors.

We found that host genetic effects were subtle but significant for the fungal, though not bacterial, root microbiome. Estimates of heritability revealed that a small number of microbial taxa showed low-to-moderate host genetic control. Interestingly, all heritable taxa were non-core members of the microbiome, and their abundance was influenced mainly by non-additive genetic effects.

These results suggest that while selective breeding may influence the abundance of certain microbial taxa, opportunities to modify whole microbiome structure are limited. Determining whether the heritable taxa affect tree fitness will be essential for evaluating the potential of microbiome-informed breeding strategies.

23. Climate Adapted Finfish: Salmon Thermotolerance and Multiomics Update

Symonds, J.E.[1], Scholtens, M.R.[1], Dodds, K.G.[2], Costilla, R.[1], Ariñez, J.B.[3], Kenny, N.J.[3], Walker, S.P.[1], Nicholson, G.[1], Waddington, Z.[4], Clarke, S.M.[2]

[1]Cawthron Institute, Nelson, NZ, [2]Bioeconomy Science Institute AgResearch, Invermay Agricultural Centre, Mosgiel, NZ, [3]Biochemistry Department, University of Otago, Dunedin, NZ, [4]The New Zealand King Salmon Co. Ltd, Picton, NZ.

The Climate Adapted Finfish programme is developing future adaptive breeding strategies for the aquaculture industry, embedded within bespoke adaptation planning pathways. One of the target species of this research is farmed Chinook salmon, as they are negatively impacted when temperatures exceed 17.0°C for prolonged periods and marine heatwaves have resulted in reduced survival. An important climate resilience trait in this species is thermotolerance. We explored the potential to improve this trait by using pedigree Chinook salmon (n=2,609) in tank-based challenges to investigate survival, time to death and growth at two temperatures: (a) 23.5 °C (HT) and (b) 21.0 °C with lower dissolved oxygen (TO). In addition, we are developing multiomic resources (genome assembly, transcriptome and methylome) to enhance genomic selection for thermotolerance.

Genetic parameters for survival, time-to-death and growth were estimated for both challenges. Analysis showed that these traits during the two challenges were moderately heritable, demonstrating the potential to improve thermotolerance through selection. However, survival and growth at TO and HT were shown to be different traits. Therefore, which challenge design is used for future commercial breeding decisions needs to be carefully considered.

At the end of both challenges, multiple tissues were collected from fish with divergent growth phenotypes for the multiomic analysis. In parallel, an annotated genome assembly for a farmed NZ Chinook salmon female from the same cohort is underway. An epigenetic clock has also been generated using reduced representation methylation sequencing from 89 fish of known age. Together, these resources will be applied for gene discovery using genome-wide association studies and to assess whether methylome profiling can be used to enhance genomic prediction. Deviations from chronological age using the epigenetic clock will be evaluated as a potential predictor of thermotolerance. These resources provide a strategic foundation for implementing climate-resilient genomic selection in Chinook salmon.

24. Development of genome assembly for climate resilience breeding of farmed Chinook salmon in New Zealand

Ariñez, J.B.[1,2,3], Symonds, J.E.[2], Scholtens, M.R.[2], Kenny, N.J.[1], Clarke, S.[3]

[1]Department of Biochemistry, University of Otago, Dunedin, NZ, [2]Cawthron Institute, Nelson, NZ, [3]AgResearch, Invermay Agricultural Center, Mosgiel, NZ.

New Zealand (NZ) is the largest producer of farmed Chinook salmon globally. However, as increased sea surface temperatures and marine heatwaves became more common in the country, mortalities in the more northerly production sites have increased. The genetic improvement of climate resilience traits serves as a promising solution to address the problem and knowledge on genomic architecture can enhance this selection. The availability of a high quality genome assembly, and its subsequent annotation, enables the identification of genetic variants that may have large effects on targeted traits. Here, we present preliminary results in assembling the genome of a female Chinook salmon which originated from one of the country's genetically distinct and locally-adapted farmed populations. Through flow cytometry and bioinformatics-based genome profiling using PacBio HiFi reads, the genome size of NZ female Chinook salmon was estimated to be 2.4-2.8 Gb. The contig N50 (2.7 Mb) and BUSCO completeness score (94.7%) were consistent and comparable to other assemblies reported by earlier studies using the same species. Further refinement, through the incorporation of ultra-long reads to resolve repetitive elements and duplicated regions in the genome, is currently being done to achieve optimum assembly results. Ultimately, the genome assembly developed in this study can be used to complement ongoing work to breed for climate-resilient Chinook salmon and support genotyping initiatives for performing quantitative genetics and comparative genomic analyses.

25. Sex prediction using genomic information and machine learning in New Zealand Sockeye salmon

Costilla, R.[1], Brauning, R. [2], Clarke S.M.[2], van Stijn T.C. [2], Wilson E. [2], Bailey J.[3], Swart J.[3], Scholtens, M.R.[1], Symonds, J.E.[1] and Dodds K.G.[2]

[1]Cawthron Institute, Nelson, NZ, [2]Invermay Agricultural Centre, Bioeconomy Research Institute, Mosgiel, NZ, [3]Mount Cook Alpine Salmon, Twizel, NZ.

Sex prediction in salmonid species is an active research area with many unanswered questions. Current evidence shows that both genetic and environmental factors affect the expression of phenotypic sex in salmonids, including Sockeye salmon (*Oncorhynchus nerka*). Sockeye salmon were successfully introduced to New Zealand over 100 years ago and wild populations remain in the South Island. However, questions remain about whether sex can be reliably predicted and the degree of discrepancies between genetic and phenotypic sex in these new environments. Here, we used genomic data from 178 samples to predict sex in New Zealand Sockeye salmon from three populations, two wild and one farmed. First, we identified "sex markers" using a *F*st test for SNPs from genotyping-by-sequencing data, e.g. SNPs with diverging allele frequencies between males and females. Sex prediction was then made using two methods for these sex markers: (1) extreme gradient boosting, a supervised learning method for prediction, and (2) a linear probability model for sex, analogous to breed prediction. Our results were promising, with the sex markers being in chromosomes LG9a and LG9b, previously associated with sex determination in Sockeye salmon. In addition, prediction rates for females were close to 100% within populations. Prediction rates for males were a bit lower ranging from 64% to 100% within populations. These results are comparable to other genetic tests for sex in Sockeye and other salmonid species. However, care might be needed when extrapolating results across populations/locations as sex prediction could be less accurate and this could delay the development of all-female broodstock for breeding programs. Overall, we demonstrate the feasibility of using genetic testing for sex in Sockeye salmon from New Zealand.

26. Genomics-Informed Breeding for Climate Resilience in Aquaculture

Leder, M.R.[1,2], Santure, A.W.[2], Wellenreuther, M.[1,2]

[1]Plant & Food Research Group, New Zealand Institute for Bioeconomy Science Ltd, Nelson, New Zealand, [2]School of Biological Sciences, University of Auckland, Auckland, New Zealand.

As marine heatwaves increase around Aotearoa New Zealand, the aquaculture industry must adapt to achieve long-term sustainability and profitability. One approach to achieve this is by diversifying its aquaculture species and developing climate-resilient breeding lines. The Australasian snapper (*Chrysophrys auratus*, tāmure) is a promising aquaculture candidate as it thrives between 18–24°C and has already achieved significant growth gains through a breeding programme established by Plant & Food Research. To further understand their growth capability and identify potential farming sites, a year-long thermotolerance trial began in April 2025 to investigate the performance of snapper across four temperature profiles. At ~5 months old, ~3,200 snapper were randomly distributed across four treatments (four tanks/treatment): an ambient profile reflecting the species' southern range in New Zealand (11.1–21.9°C), a warm and cold profile set 2.5°C above and below ambient, respectively, and a stable optimal profile mimicking recirculating aquaculture systems at 21°C. Mortalities are recorded daily, and weight and length measurements are taken for all individuals every eight weeks. To identify climate-resilient genotypes, growth and survival data will be linked to genotyping-by-sequencing data for all fish. These genetic and phenotypic datasets will be integrated to analyse genotype-by-environment interactions, pinpoint quantitative trait loci that influence performance, reconstruct family pedigrees, and calculate the heritability of resilience traits. Four months into the trial, cumulative specific growth rate (% weight gain/day, mean ± SD) was 0.97 ± 0.13, 0.39 ± 0.10, 0.17 ± 0.08, and 0.04 ± 0.05 in the optimal, warm, ambient, and cold profiles, respectively. Approximately six months into the trial, overall mortality was low at 2.6%. The genomics-based selection, to be performed at the end of the experiment, will help inform future selective breeding programmes to enhance the resilience of cultured snapper stocks to temperature variability around New Zealand.

27. Tools to accelerate genetic gain for *Ostreid Herpesvirus-1* (OsHV-1) resilience in Pacific oysters (*Crassostrea gigas*)

Scholtens, M.R.[1], Rolton, A.[1], Delisle, L.[1], King, N.[1], Vignier, J.[1]

[1]Aquaculture Group, Cawthron Institute, Private Bag 2, Nelson 7042, NZ.

Selective breeding has delivered measurable improvements in *Ostreid herpesvirus-1* (OsHV-1) resilience in Pacific oysters (*Crassostrea gigas*), yet reliance on costly challenge tests and destructive DNA sampling methods limits the efficiency of breeding programs. We therefore explored two potential tools to address these constraints: haemolymph traits as indicator phenotypes of OsHV-1 survival, and swabs as a non-lethal alternative to gill excision for genotyping. Fifteen haemolymph traits were analysed in nine full-sib families representing divergent survival backgrounds. Bivariate animal models (>180,000 pedigree records) were fitted to estimate heritabilities and genetic correlations (rG) with survival under both field and laboratory challenges. For DNA sampling, paired swab and gill samples (62 oysters; 124 total) were genotyped using genotyping-by-sequencing and compared for data quality, genotype concordance, minor allele frequency, and pedigree applications.

Four haemolymph traits showed strong genetic correlations with survival. The reactive oxygen species production of the entire haemolymph (ROSHem) and of the hyalinocyte sub-population (ROSHya) were negatively associated with both field (rG = –0.92 ± 0.34 and -0.96 ± 0.32, respectively) and laboratory survival (rG = –0.62 ± 0.29 and -0.63 ± 0.28, respectively), while the lysosomal content of hyalinocytes (LysomeanHya) and entire haemolymph (LysomeanHem) were positively correlated with laboratory survival (rG = 0.82 ± 0.21 and 0.90 ± 0.17, respectively). These haemolymph traits were moderately heritable (up to 0.58 ± 0.25). Swab-derived genotypes were highly concordant with the gill samples and produced equivalent estimates of relatedness, inbreeding, and parentage assignment.

These findings demonstrate that haemolymph traits and swab sampling provide practical tools to reduce reliance on destructive assays, improve operational efficiency, and accelerate genetic gain in oyster breeding programs.