XIX Conference 2025

EUCARPIA

BIOMETRICS IN PLANT BREEDING

17-19 Sep 2025 Edinburgh, UK

BOOK OF ABSTRACTS

The EUCARPIA Biometrics in Plant Breeding
Local Organising Committee

This is a short version of the abstract booklet for print use.
Full details on the conference, including social events and conference dinner, can be found at:
`https://highlanderlab.github.io/EUCARPIA2025BiometricsPlantBreeding/`

Contents

Welcome **5**
- EUCARPIA Biometrics Section Chair . . . 5
- Local Organising Committee . . . 6

About **7**
- EUCARPIA Biometrics in Plant Breeding, Edinburgh 2025 . . . 7
- Local Organizing Committee . . . 7
- International Scientific Board . . . 7

Sponsors **8**
  - Platinum sponsors . . . 8
  - Gold sponsors . . . 9
  - Silver sponsors . . . 9

Invited Speakers **10**
- Gary Atlin — Bill & Melinda Gates Foundation, USA . . . 10
- Rosemary Bailey — University of St Andrews, Scotland . . . 10
- Hao Cheng — University of California, Davis, USA . . . 11
- Keith Gardner — CIMMYT, Mexico . . . 11
- Steven Penfield — John Innes Centre, England . . . 12
- María Xosé Rodríguez-Álvarez — University of Vigo, Spain . . . 12
- Pascal Schopp — KWS Group, Germany . . . 12
- Julian Taylor — University of Adelaide, Australia . . . 13
- Andrea Doeschl-Wilson — The Roslin Institute, Scotland . . . 13

Timetable **14**
- Wednesday, 17 September . . . 14
- Thursday, 18 September . . . 16
- Friday, 19 September . . . 17

Wednesday 17th September **19**
- Session 1: 9:00-10:40 . . . 19
- Session 2: 11:10-12:50 . . . 25
- Session 3: 13:50-15:30 . . . 31
- Session 4: 16:00-17:40 . . . 37
- Poster session 1 & Welcome reception: 17:40-19:40 . . . 43

Thursday 18th September **45**
- Session 5: 9:00-10:40 . . . 45
- Session 6: 11:10-12:50 . . . 51
- Social events & Conference dinner: 13:20-23:30 . . . 57

Friday 19th September **59**
- Session 7: 9:00-10:40am . . . 59

Session 8: 11:10-12:50 65
Session 9: 13:50-15:30 71
Poster session 2 & Farewell reception: 15:30-17:00 77

Poster session 1 **78**

Poster session 2 **114**

Welcome

EUCARPIA Biometrics Section Chair

A word of welcome

It is an honour to welcome you to the 19th meeting of the EUCARPIA Section “Biometrics in Plant Breeding”. Like the other sections of EUCARPIA, our section aims at facilitating interactions and discussions between researchers in plant breeding. For the Biometrics in Plant Breeding section, the topics are related to the use of quantitative methods to improve breeding and selection. I am very pleased that the scientific committee has been able to compose an excellent programme comprising talks and posters spanning nine specific topics, each lead off by one of our invited speakers.

The scientific programme as a whole comprises nine invited speakers, 36 contributed speakers and more than 70 poster presenters. There will be ample opportunity to exchange ideas all along the conference, but in this context I would like to recommend especially the two poster sessions, where refreshments will be served as well. The social programme provides a choice between several excursions on Thursday afternoon and a conference dinner on Thursday evening. I wish to thank the local and scientific committees for the organization and the sponsors for extra financial support. I am also very happy that the invited speakers were willing to accept the invitations of the scientific committee and that so many researchers made the effort to submit contributed papers and posters. On behalf of the International Scientific Committee I hope you will enjoy the meeting.

Hans-Peter Piepho
Chairman of the Biometrics in Plant Breeding Section of EUCARPIA

Local Organising Committee

On behalf of the Local Organising Committee, it is our great pleasure to welcome you to the XIX EUCARPIA Biometrics in Plant Breeding Conference, Edinburgh 2025.

The EUCARPIA Biometrics Conference has a long tradition of bringing together researchers, breeders and industry to exchange ideas and strengthen collaborations in the field of biometrics applied to plant breeding. We are especially delighted to host this year's meeting in Edinburgh, a city with a distinguished history in genetics, statistics and agriculture. From the influential contributions of the Edinburgh School of Genetics and Biomathematics in the 20th century, to the continuing advances in biometrics and quantitative genetics at the Roslin Institute, Biomathematics & Statistics Scotland (BioSS) and Scotland's Rural College (SRUC), Edinburgh has long been at the forefront of developing and applying quantitative methods in breeding and genetics. Today, the city remains an international hub for research and collaboration, and we hope that this scientific heritage, together with Edinburgh's vibrant cultural setting, will inspire both fruitful discussions and enjoyable experiences beyond the conference halls.

Organising this conference happened somewhat by chance, as many things in life do. Ian MacKay invited us to support his bid to host this eminent meeting, but sadly he was not able to see it through, and so the responsibility of organisation fell to us. While unexpected, we are fortunate to have had the chance to bring the conference to Edinburgh and feel privileged to honour Ian's legacy. To recognise his lasting contributions to biometrics and quantitative genetics in plant breeding, we are pleased to introduce an award for the best presentation and poster from early career researchers. In this way, we hope to celebrate Ian's influence by encouraging and recognising the next generation of scientists in our field. Such initiatives, together with the wider conference activities, have been made possible through the generous support of EUCARPIA and our industry partners (Platinum: BASF, VSNi; Gold: AbacusBio, Bayer CropSciences, GDM, Rijk Zwan; Silver: Driscoll's, KWS, NPZ).

We sincerely thank the invited speakers and special guests for accepting our invitation, those whom have contributed talks and posters, and all participants for joining us. We hope you enjoy the programme, meet new and old colleagues, see a bit of Edinburgh, and have fun!

Daniel Tolhurst and Gregor Gorjanc
Leads of the Local Organising Committee

About

EUCARPIA Biometrics in Plant Breeding, Edinburgh 2025

The EUCARPIA Biometrics in Plant Breeding Conference serves as a premier platform for researchers, breeders and industry to explore cutting-edge methods in biometrics, statistical genetics and practical plant breeding. The scientific programme features nine thematic sessions, each anchored by a plenary presentation followed by four contributed talks. Highlights include discussions on practical implementation of biometrics in public and private programmes, design of experiments and multi-environment trials, genetic and genomic modelling, leveraging genotype by environment interaction, high-throughput phenotyping, artificial intelligence and machine learning applications and strategies for managing genetic diversity to maximise long-term gains. Together, the programme reflects both the methodological foundations of biometrics and their real-world application to the challenges facing modern plant breeding.

The conference will take place from 17-19 September at the Royal College of Physicians on Queen Street, Edinburgh. The pre-conference AlphaSimR and FieldSimR workshop will be held on 15-16 September at the Roslin Institute, Easter Bush, Midlothian. On Thursday afternoon, a programme of social events will showcase different aspects of Edinburgh, followed by the conference dinner at Mansfield Traquair on Broughton Street – a stunning hidden gem in the heart of the city. We look forward to a lively and engaging conference and sharing an unforgettable dinner with you, complete with nice touches of Scottish history and traditional Scottish dancing.

This is a short version of the programme booklet for print use. Full details on the conference, including social events and conference dinner, can be found at:
`https://highlanderlab.github.io/EUCARPIA2025BiometricsPlantBreeding/`

Local Organizing Committee

Daniel Tolhurst
Jaime Ortiz
Hannes Becher
Gabriela Mafra Fortuna
Gregor Gorjanc
Yu Zhang
Duncan Henderson
Isidore Houaga
Rosalind Craddock
Dominic Waters
Ezekiel Okeleye

International Scientific Board

Hans-Peter Piepho
Daniel Bustos-Korts
Chris Maliepaard
Laurence Moreau
Daniel Tolhurst
Lucia Gutierrez
Tristan Mary-Huard
Guillaume Ramstein
Gregor Gorjanc
Pawel Krajewski
Emilie Millet
Valentin Wimmer

Sponsors

Platinum sponsors

BASF

We create chemistry

EUCARPIA

European Association for Research on Plant Breeding
Europäische Gesellschaft für Züchtungsforschung
Association Européenne pour l’Amélioration des Plantes

VSNi

Gold sponsors

abacusbio

BAYER

GDM

RIJK ZWAAN

Silver sponsors

Driscoll's
Only the Finest Berries™

SEEDING
THE FUTURE
SINCE 1856

KWS

NORDDEUTSCHE PFLANZENZUCHT
Wir lassen Qualität wachsen.

NPZ

Invited Speakers

Gary Atlin — Bill & Melinda Gates Foundation, USA

Dr Gary Atlin received BS and MS degrees in crop science at the University of Guelph, then worked as a research associate in the potato program at CIP in Peru for two years. His PhD work in oat breeding at Iowa State University focused on the design of breeding programs for stressful environments. He worked as a commercial flax breeder in Alberta, Canada, then taught plant breeding at Nova Scotia Agricultural College for ten years, developing a theoretical framework for managing components of genotype x environment interaction in breeding pipelines in order to maximize genetic gain. In 2000, Gary joined IRRI as a rice breeder, establishing IRRI's drought tolerance pipeline, which resulted in the development of rice varieties with improved drought tolerance that were widely grown in South Asia. His group also identified the first large-effect QTLs affecting rice yield under drought stress. He joined CIMMYT in 2006 as a maize breeder, and became technical breeding lead in 2009, coordinating efforts to optimize CIMMYT's maize pipelines to increase genetic gains. In 2012, he joined the Gates Foundation as a Senior Program Officer in the Agricultural Development R&D team. He currently manages the foundation's maize, wheat, and rice breeding portfolios. Gary has helped design and implement strategies to increase the rate of genetic gain delivered to smallholders in Sub-Saharan Africa and South Asia. Recently, Gary has worked with CGIAR colleagues on the implementation of genomic selection in breeding networks serving highly diverse African cropping systems, using innovative designs to estimate the breeding value of selection candidates dispersed across research stations and farms sampling the target population of environments.

Rosemary Bailey — University of St Andrews, Scotland

Rosemary Bailey is Professor of Statistics at the University of St Andrews, and a Fellow of the Royal Society of Edinburgh. Rosemary worked for the Medical Research Council's Air Pollution Research Unit before studying at the University of Oxford, where she obtained a BA in Mathematics and a DPhil in Group Theory. As a post-doctoral fellow at the University of Edinburgh, Rosemary learnt how to apply group theory to problems in design of experiments. Rosemary then spent ten years applying this knowledge in the Statistics Department at Rothamsted Experimental Station, before moving to academia as Head of Department or School at Goldsmiths College and at Queen Mary College, both in University of London. Rosemary was President of the then-British Region of the International Biometric Society from 2000 to 2002, and has also served on various committees of the London Mathematical Society, the Royal Statistical Society and the Institute of Mathematical Statistics.

Hao Cheng — University of California, Davis, USA

Professor Hao Cheng's research is broadly involved in the development of statistical, machine learning, and computational methods that bridge the genome and phenome for the genetic improvement of populations through more accurate, efficient, and biologically meaningful analyses. Hao has focused on the use of genomics, phenomics, and other sources of omics data in various species to better predict desired traits and infer the mechanisms underlying them. This includes the theoretical aspects of quantitative genetics, such as statistical models and computational algorithms (e.g., genomic prediction and association studies), the development of software tools to apply these statistical methods and computational algorithms to real-life, large-scale omic data, and the application of these methods and tools to large datasets for a wide range of traits in various species.

Keith Gardner — CIMMYT, Mexico

Dr Keith Gardner is head of the Biometrics and Quantitative Genetics Unit at the International Maize and Wheat Improvement Center, CIMMYT. CIMMYT is a non-profit, international applied agricultural science organization dedicated to fostering improved quantity, quality, and dependability of production systems and basic cereals such as maize, wheat, triticale, sorghum, millets, and associated crops, particularly in the Global South. Through this endeavour, CIMMYT enhances the livelihood trajectories and resilience of millions of resource-poor farmers, while working towards a more productive, inclusive, and resilient agrifood system within planetary boundaries. As head of the Biometrics and Quantitative Genetics Unit, Keith works with CIMMYT breeding teams and national partners to develop efficient analysis pipelines, to optimize breeding schemes to accelerate genetic gain, and to effectively manage the deployment of novel useful variation from discovery research into plant breeding programs. Keith holds a PhD in Evolutionary Biology and Genetics from Indiana University and prior to working at CIMMYT, he was Program Leader in Quantitative Genetics at NIAB, a non-profit applied agricultural research institute based in Cambridge, UK.

Steven Penfield — John Innes Centre, England

Professor Steven Penfield's research group works to understand how weather and climate affect plant reproductive development and seed quality. Steven is particularly interested in the effects of temperature on crop yields, seed quality and seedling establishment. Steven's work uses crop species from the Brassica family, namely oilseed rape (canola) and vegetable Brassicas, as well as model species such as *Arabidopsis thaliana*. In this way, he aims to understand precise mechanisms by which weather and climate affect crop performance, and consider strategies for increasing crop resilience, for instance using conventional plant breeding or genome editing techniques.

María Xosé Rodríguez-Álvarez — University of Vigo, Spain

Dr María Xosé Rodríguez Álvarez earned her PhD in Mathematics from the Universidade de Santiago de Compostela, Spain, in 2011 and has a diverse professional background spanning both the private sector and academia. Since 2021, María has been a Ramón y Cajal fellow at the Universidade de Vigo. Her research focuses on (i) developing efficient estimation methods for flexible regression models, (ii) statistically evaluating the diagnostic and prognostic value of clinical biomarkers, and (iii) proposing new statistical methods for analysing spatial and spatio-temporal processes in the context of agricultural experiments. Her work emphasises practical applications and interdisciplinary collaboration, with a strong commitment to disseminating advancements through free software.

Pascal Schopp — KWS Group, Germany

Dr Pascal Schopp is the Head of Corn Breeding Europe Mid-Late at KWS, where he is responsible for germplasm and product development in the mid-early to mid-late dent grain corn markets. He leads multiple breeding programs across Europe. Pascal holds an MSc in Plant Breeding from the University of Hohenheim, Germany, where he also completed his PhD at the Institute of Applied Plant Breeding. During his PhD studies, Pascal worked on genomic prediction methodology and its application in plant breeding.

Julian Taylor — University of Adelaide, Australia

Julian Taylor is an Associate Professor in the Biometry Hub at the Waite Agricultural Precinct of the University of Adelaide (UA). Julian is the Node Leader of the Analytics for the Australian Grains Industry in University of Adelaide (AAGI-UA), Grains Research and Development Corporation (GRDC) funded initiative to provide national analytics research support for the Australian Grains Industry. Within AAGI-AU, Julian strategically manages a multi-disciplinary analytics teams across four schools and fosters the development of collaborative high impact analytics research, support and training activities. AAGI-AU, and other partners on the network, have broadened the analytics toolbox within the Australian grains industry to include biometry, machine learning, artificial intelligence, data science, mathematics as well as computational infrastructure to support the breadth of these analytics activities.

Andrea Doeschl-Wilson — The Roslin Institute, Scotland

Andrea Doeschl-Wilson is Professor of Animal Disease Genetics and Modelling at the Roslin Institute of the University of Edinburgh in the UK, where she leads a research group. Her research group uses mathematical modelling to assess and predict how genetic and non-genetic factors influence host responses and impact to the infectious or harmful social environment of farm animals. Andrea also leads the Roslin Institute Strategic Programme on the Prevention and Control of Infectious Diseases.

Timetable

IS: Invited Speaker CT: Contributed Talk

Wednesday, 17 September

8:00–8:30	**Registration & tea/coffee**		
8:30–9:00	**Welcome: Daniel Tolhurst & Gregor Gorjanc**		
9:00–10:40	**Session 1: Impact of biometrics in public and private breeding programmes**		
9:00–9:40	IS	**Gary Atlin**	Early stage sparse testing can increase selection accuracy and genetic gain in plant breeding programmes
9:40–9:55	CT	**Christian Werner**	Reviving the Desired Gains Index: an optimal solution for parent selection in plant breeding programmes
9:55–10:10	CT	**Wubishet Bekele**	Lessons from a decade of streamlining genomic selection in oat breeding programs at AAFCs
10:10–10:25	CT	**Augustin Desprez**	Twenty years of asymmetrical hybrid breeding drives diversity and targets domestication genes in commercial elite sugar beet inbreds
10:25–10:40	CT	**Leticia Lara**	Trait prioritization and economic modelling to guide strategic decisions in pea breeding program
10:40–11:10	**Morning tea/coffee**		
11:10–12:50	**Session 2: Design of experiments and multi-environment trials**		
11:10–11:50	IS	**Rosemary Bailey**	Designs with more than one blocking system: the conflict between valid randomization and latinization
11:50–12:05	CT	**Emi Tanaka**	Leveraging large language models in a composable system to design and simulate plant experiments
12:05–12:20	CT	**Karen Wolf**	On deficient experimental designs and their analysis
12:20–12:35	CT	**Michaela Jung**	Designing multi-environment reference populations in perennials: do’s and don’ts
12:35–12:50	CT	**Bjarne Nielsen**	Power calculations for testing genetic correlation between pure- and mixture line effects in wheat breeding
12:50–13:50	**Lunch**		

13:50–15:30	**Session 3: Genetic and genomic modelling of plant breeding data**		
13:50–14:30	IS	**Pascal Schopp**	A genomic selection reality check from industry: what moves the needle in genetic gain
14:30–14:45	CT	**Yong Jiang**	An efficient and powerful algorithm for detecting heterotic QTL
14:45–15:00	CT	**Tristan Mary-Huard**	A composite hypothesis testing approach to detect pleiotropic genomic regions
15:00–15:15	CT	**Antero Heikkilä**	Approximating prediction error variances of genomic models using Monte Carlo sampling
15:15–15:30	CT	**Javier Fernández-González**	Why REML variance estimates have no biological meaning and how to solve it
15:30–16:00	**Afternoon tea/coffee**		
16:00–17:40	**Session 4: Innovative models with unconventional data streams**		
16:00-16:40	IS	**Hao Cheng**	Beyond mixed models: integrative models for agricultural G2P using unconventional data streams
16:40-16:55	CT	**Hannah Robinson**	Leveraging multi-omics for predictive modelling in perennial crops: a case study in grapevine
16:55-17:10	CT	**Aiyesa Leke Victor**	Genomic prediction at the root: from one plant to better populations
17:10-17:25	CT	**Timothée Flutre**	Quantifying the genetic basis of mixing ability and investigating its genetic architecture: case studies in wheat varietal mixtures and wheat-pea intercrops
17:25-17:40	CT	**Nicolas Salas**	When durum wheat grows together with alfalfa: The Genetic Dialogue of Neighboring Plants
17:40–19:40	**Poster session 1 & Welcome reception**		

Thursday, 18 September

8:00–8:50	**Tea/coffee**		
8:50–9:00	**Welcome**		
9:00–10:40	**Session 5: Genotype performance, stability and resilience**		
9:00-9:40	IS	**Andrea Doeschl-Wilson**	G x infectious E: breeding for reduced pathogen transmission
9:40-9:55	CT	**Emilia Koch**	A hurdle model for ordinal scoring data with an underlying percentage scale
9:55–10:10	CT	**Lucia Gutierrez**	How many environments do we really need for yield stability prediction in cereals
10:10–10:25	CT	**Dominic Waters**	The role of stability metrics in plant breeding programmes
10:25–10:40	CT	**Hugo Dorado**	Estimating genetic performance and variability on-farm using ranking data
10:40–11:10	**Morning tea/coffee**		
11:10–12:50	**Session 6: Modelling and leveraging genotype by environment interaction**		
11:10–11:50	IS	**Steven Penfield**	Exploiting genotype by environment interactions to breed for local adaptation and climate change resilience in winter rapeseed
11:50–12:05	CT	**Shanice Van Haeften**	Characterising genotypic stability and environmental responsiveness in Australian chickpea
12:05–12:20	CT	**Jip Ramakers**	Multiple-covariate random regression for predicting yield across the Australian wheatbelt
12:20–12:35	CT	**Killian Melsen**	REML implementations of kernel-based multi-trait, multi-environment genomic prediction models
12:35–12:50	CT	**Wera Eckhoff**	Tailoring AI and ML models for genotype-by-environment prediction leveraging environmental covariates: a European rye example
12:50–13:20	**Packed lunch**		
13:20–17:00	**Social events**		
18:00–23:30	**Conference dinner at Mansfield Traquair**		

Friday, 19 September

8:00–8:50	**Tea/coffee**		
8:50–9:00	**Welcome**		
9:00–10:40	**Session 7: High-throughput phenotyping**		
9:00-9:40	IS	**María Xosé Rodríguez-Álvarez**	Spatio-temporal modelling of genotypic performance in high-throughput phenotyping
9:40-9:55	CT	**Corné Verburg**	Bridging genomic prediction and crop growth modeling with System Identification for Nonlinear Dynamics (SINDy)
9:55–10:10	CT	**Timothy Thavarajah**	Deploying computer vision and bivariate genomic prediction to improve blackleg disease resistance in canola breeding
10:10–10:25	CT	**Ángela Prudencio**	Integration of new RFID-based pollinator parameters into plant phenotyping and breeding selection indexes under climate change pressure
10:25–10:40	CT	**Clément Bienvenu**	Optimizing GxE prediction in phenomic selection using chemometric decomposition of NIRS Data
10:40–11:10	**Morning tea/coffee**		
11:10–12:50	**Session 8: Artificial intelligence and machine learning in practice**		
11:10–11:50	IS	**Julian Taylor**	Integrating artificial intelligence and machine learning capability into Australian grains agriculture
11:50–12:05	CT	**Tally Wright**	Exploring random forest-based genomic prediction for optimised selection of reference barley varieties for inclusion in DUS distinctiveness evaluation
12:05–12:20	CT	**Carles Quesada-Traver**	Opportunities and limitations of deep learning-based multi-environment genomic prediction
12:20–12:35	CT	**Natasha Johansen**	Evolutionary-scale protein language models effectively uncover beneficial variants in a sorghum bicolor diversity panel
12:35–12:50	CT	**Sarah Ghysels**	Image-based yield prediction for tall fescue using random forests and convolutional neural networks
12:50–13:50	**Lunch**		

13:50–15:30	**Session 9: Managing genetic diversity and maximising long-term genetic gains**		
13:50–14:30	IS	**Keith Gardner**	Effective and efficient utilization of gene bank resources using quantitative genetics
14:30–14:45	CT	**Carolina Rivera-Poulsen**	A rapid cycling genomic selection experiment in maize landraces
14:45–15:00	CT	**Kira Villiers**	An algorithm for mate selection by haplotype stacking
15:00–15:15	CT	**Sónia Surgy**	The maximisation of the genetic gains of the polyclonal selection in grapevine varieties using integer programming
15:15–15:30	CT	**Seifelden Metwally**	A novel framework to control genetic diversity for optimal genomic mating
15:30–17:00	**Poster session 2 & Farewell reception**		
17:00	**Conference end**		

Wednesday 17th September

Session 1: 9:00-10:40

Impact of biometrics in public and private breeding programmes

Chair: Wayne Powell – SRUC, Scotland

Early stage sparse testing can increase selection accuracy and genetic gain in plant breeding programmes

Gary Atlin[1], *David Gonzalez Dieguez*[2,3], *Juan Burgeno*[3], *Teshale Mamo*[4], *Yoseph Beyene*[5], *Dagne Wegary*[6], *Jill Cairns*[6], *Walter Chivasa*[5], *Mainassarah Zaman-Allah*[6], *Christian Werner*[2,3] IS

[1]Bill & Melinda Gates Foundation, Seattle, Washington, USA
[2]Accelerated Breeding (B4T), Consultative Group of International Agricultural Research (CGIAR), Texcoco, Mexico
[3]International Maize and Wheat Improvement Center (CIMMYT), Texcoco, Mexico
[4]Alliance of Bioversity-CIAT, Arusha, Tanzania
[5]CIMMYT, Nairobi, Kenya
[6]CIMMYT, Harare, Zimbabwe

Low-cost genome profiling allows information from relatives to be used to reduce replication of selection candidates while retaining selection accuracy, creating opportunities to redesign plant breeding pipelines. Early-stage sparse testing uses the genomic relationship matrix (GRM) to sample the target population of environments (TPE) more effectively by distributing related early-stage selection candidates across many testing locations to train a genomic selection (GS) model, rather than concentrating them in replicated trials at a few research stations. Using the GRM, data on related genotypes across farms can be connected and combined to enable genomic prediction across the entire TPE, not just a few sites. A similar redesign is possible in hybrid breeding programmes, where the GRM allows related selection candidates to be evaluated on several testers, permitting general combining ability (GCA) to be estimated early in the selection process. Simulations indicate that, if testing resources are held constant, sparse testing designs can increase selection accuracy both across environments and across testers.

A promising application of sparse testing is to move early-stage testing directly into farmers' fields. Most crop production in Sub-Saharan Africa (SSA) occurs on small farms characterized by low input use, multiple biotic and abiotic stresses, and diverse management factors. However, most CGIAR and national breeding programmes conduct early-stage phenotyping at only a few well-managed research stations. Early-stage on-farm sparse testing (OFST) addresses this issue by shifting the evaluation process from research stations to dozens or even hundreds of farms, treating each farm as an incomplete block. This can enhance genetic gain in farmers' fields by increasing selection accuracy and intensity while shortening the generation interval. Early-stage OFST is currently being piloted in the CIMMYT maize and CIAT bean programmes in SSA.

Reviving the Desired Gains Index: an optimal solution for parent selection in plant breeding programmes

Christian Werner[1], *Daniel Tolhurst*[2]

[1] Accelerated Breeding (B4T), Consultative Group of International Agricultural Research (CGIAR), Texcoco, Mexico
[2] The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, United Kingdom

The Desired Gains Index (DGI) offers an optimal solution for parent selection in public plant breeding programmes and small to medium-sized breeding enterprises. It allows breeders to express breeding objectives as desired responses to selection, thereby supporting simple and efficient multi-trait improvement.

In this presentation, we challenge the widespread perception that the DGI is less efficient than profit-driven indices popular in animal breeding, such as the Smith-Hazel Index. We deliberately refer to the DGI as "optimal" to emphasise that, like profit-based indices, it maximises genetic gain relative to the breeding objective. However, deriving accurate economic weights for profit-driven indices is often too complex, costly, and impractical in breeding programmes lacking the resources to link traits directly to profitability. Moreover, the relationship between trait improvement and profit can be nonlinear, especially in traits not directly tied to yield or market value. Conversely, the DGI offers an intuitive approach to derive breeding objectives directly from a target variety concept.

First, we review how desired gains inherently represent a form of economic weighting, expressed as trait improvement ratios. This demonstrates that the DGI and the Smith-Hazel Index are not mutually exclusive but offer complementary perspectives on multi-trait selection. Then, we show how understanding this relationship enables the derivation of economic weights from desired gains, and vice versa. Thereby, breeders gain access to an optimisation framework which allows them to evaluate and refine their breeding objective in a reciprocal manner. The DGI remains the preferred tool for multi-trait selection in CGIAR plant breeding programmes.

Lessons from a decade of streamlining genomic selection in oat breeding programs at AAFC

Wubishet Bekele[1], *Weikai Yan*[1,2], *Nicholas Tinker*[1], *Mehri-Hadinezhad*[1], *Jennifer Mitchel-Fetch*[2], *Judith Frigeau-Reid*[1], *Asuka Itaya*[1], *Brad DeHaan*[1], *Matthew Hayes*[1], *Kali Stewart*[2], *Savka Orozovic*[1], *Charlene Wight*[1], *Dan MacEachern*[3], *Aaron Beattie*[4], *Nathan Mountain*[6], *Melinda Drummond*[6], *Allan Cummiskey*[3], *Denis Pageau*[5], *Isabelle Morasse*[5], *Geneviève Telmosse*[5], *Holly Byker*[6], *Helen Booker*[6], *Mike Holzworth*[6]

[1]Ottawa Research and Development Centre, Agriculture and Agri-Food Canada
[2]Brandon Research and Development Centre, Agriculture and Agri-Food Canada
[3]Charlottetown Research and Development Centre, Agriculture and Agri-Food Canada
[4]Department of Plant Sciences, Crop Development Centre, University of Saskatchewan
[5]Quebec Research and Development Centre, Agriculture and Agri-Food Canada
[6]Department of Plant Agriculture, University of Guelph

Over the past ten years, the Oat Breeding and Genomics team at Agriculture and Agri-Food Canada (AAFC) has developed and refined an integrated pipeline for genomic selection, encompassing field, greenhouse, laboratory, bioinformatics, and biometrics workflows. This streamlined system currently supports the annual evaluation of up to 10,000 oat lines, delivering genomic estimated breeding values (GEBVs) in parallel with conventional visual selection.

To date, the program has amassed genotypic and phenotypic data for over 50,000 training and test lines. This presentation will share key insights from this decade-long effort, including:

- The implementation of a cost-effective genotyping protocol that reduced per-sample costs from over CAD 35 to under CAD 15;
- Development and deployment of mega-environment-specific genomic prediction models;
- Insights gained from over six years of direct comparisons between genomic selection (GS) and visual selection (VS) strategies, highlighting the effectiveness of GS to select yield per se, and the integrative power of VS for yield, agronomic, and quality;
- Case studies of two oat varieties resulted from the combined use of VS and GS.
- Outlook: The programs are now exploring genomic selection of parental combinations and rapid-cycle recurrent selection for both spring and winter oat germplasm, aiming to accelerate genetic gain and further enhance breeding efficiency.

Twenty years of asymmetrical hybrid breeding drives diversity and targets domestication genes in commercial elite sugar beet inbreds

Augustin Desprez[1,2], *Karine Henry*[3], *Bruno Desprez*[2], *Pierre Devaux*[2], *Alain Charcosset*[1], *Maud Tenaillon*[1], *Laurence Moreau*[1]

[1]GQE - Le Moulon, INRAE, CNRS, AgroParisTech, Université Paris-Saclay, Gif-sur-Yvette, France
[2]Florimond Desprez Veuve & Fils SAS, Cappelle-en-Pévèle, France
[3]United Beet Seeds (UBS), Tienen, Belgium

Hybrid breeding in sugar beet (*Beta vulgaris* ssp. *vulgaris*) relies on a three-way cross system leveraging cytoplasmic male sterility (CMS) to produce uniform, high-yielding varieties through the coordinated use of genetically distinct male and female parental pools. Using a dataset of 11,099 plants across 1,337 accessions genotyped with high-quality SNP arrays, we investigated temporal genomic patterns of diversity, differentiation, and selection over 21 years of breeding within a major European program. Population structure analyses using PCA and ADMIXTURE revealed strong and stable differentiation between male and female pools, primarily driven by directional selection in the pollinator pool and historical founder effects. Genetic diversity within pollinators decreased significantly over time, while differentiation between heterotic pools, as measured by Fst, steadily increased, reflecting enhanced allelic complementarity. Genome-wide selection scans identified a major differentiation signal on chromosome 2 overlapping with the bvBTC1 locus, a key domestication gene regulating vernalization-dependent flowering. This region exhibited significant outlier patterns in both Fst and PCAdapt analyses, suggesting recent heterotic-pool-specific selection pressure. The asymmetrical structure of CMS-based breeding, characterized by differences in breeding methods and bolting conditions between male and female lines, likely imposed divergent selection around bvBTC1, potentially contributing to shifts in reproductive characteristics regarding hybrid seed production. This study provides novel insight into how breeding strategies can shape genetic diversity structuring and genomic patterns at key agronomic loci, even over short timescales. A similar framework could be applied to examine similar dynamics in other hybrid crop systems.

Trait prioritization and economic modelling to guide strategic decisions in pea breeding program

Leticia A. de C. Lara[1], *Tom Kirk*[1], *Mason Chizk*[1], *Natalie Howes*[2], *Martin Howes*[2], *John Crowley*[2], *Jagroop Kahlon*[3], *Jenn Walker*[3], *Leanne Fischbuch*[3]

[1]AbacusBio International, Roslin Innovation Centre Easter Bush, Edinburgh, United Kingdom
[2]AbacusBio Canada, Agri-Food Discovery Place, Edmonton, Alberta, Canada
[3]Alberta Pulse Growers, Leduc, Alberta, Canada

To encourage grower-supported breeding initiatives from Alberta Pulse Growers (APG), AbacusBio conducted a simulation and economic analysis of a pea breeding program in Western Canada. Using AlphaSimR, we modelled genetic and economic outcomes of four long-term selection strategies over a 20-year horizon, targeting yield, seed protein content (SPC), and fusarium disease resistance (FDR). The baseline scenario reflected traditional selection based on yield with constraints on SPC and FDR. Three alternative scenarios implemented economic selection indices to optimize trait combinations based on predicted grower profitability. Our results demonstrate that economic selection indices can significantly enhance both genetic gain and return on investment (ROI). Here, ROI refers to the economic benefit generated from a breeding program investment that includes genotyping, phenotyping, field trials, and program management – without increasing operational costs relative to the baseline. The most balanced scenario – optimizing for yield and FDR while maintaining a minimum SPC threshold – more than doubled the economic benefit compared to the baseline, increasing net present value from $267M to $712M and ROI from 14.1% to 19.3%. A scenario prioritizing SPC improvement produced the highest total benefit ($1,097M), however, this depended on future market compensation for SPC. Without any premium applied for SPC, the benefits were reduced to $392M. Excluding SPC from selection resulted in the highest gains in yield and FDR, with $826M in benefits but with drastic reduction in SPC. This work highlights how trait prioritization and economic modelling can guide strategic breeding decisions and maximize long-term value.

Session 2: 11:10-12:50

Design of experiments and multi-environment trials

Chair: Sue Welham – Stats4Biol, England

Designs with more than one blocking system: the conflict between valid randomization and latinization

Rosemary Bailey

IS

University of St Andrews, United Kingdom

In experimental designs with more than one blocking system, there are (at least) two different approaches. One goes back to the work of Frank Yates and John Nelder, who both worked at Rothamsted Experimental Station (but at different times). They took the view that if Rows and Columns are two necessary blocking systems, and each intersection of a Row and a Column is called a Cell, then Cells must also be included as a blocking system. Moreover, they did not regard the variability between different plots as given by a simple equation. They said that a method of randomization was valid if, when averaged over the outcomes of all possible randomizations, the expectation of the mean square for a given treatment effect is the same as the expectation of the mean square for the residual that it will be compared with, in the case that that treatment effect is zero.

The other approach, called latinization, seems to have begun with Boyd Harshbarger and Lyle Davis. The linear model is written in the standard form, with some factors having fixed effects and others random effects. There is no restriction on the relationship between pairs of blocking systems. However, in each system, the number of times that a treatment occurs in any block differs by no more than one from the number times that any other treatment occurs in that block.

I will try to explain the two approaches, and then give examples where they differ.

Leveraging large language models in a composable system to design and simulate plant experiments

Emi Tanaka

Australian National University, Australia

Designing and simulating plant breeding experiments demands careful coordination of genetic, environmental, and management factors. We present a composable system that leverages large language models (LLMs) to streamline the end-to-end workflow for designing and simulating plant experiments. Through a modular, composable interface built on the “edibble” R package, experimenters can assemble, modify, and reuse experimental components enabling rapid prototyping, customization, and replication of complex experimental designs. The extended system, built on a LLM, automates simulation of breeding outcomes, leverages historical data and information, and provides context to support informed experimental design decisions. This system empowers breeders to design robust experiments, explore potential genetic gains, and accelerate cultivar development by enhancing rigor, adaptability, and discovery in plant breeding research.

On deficient experimental designs and their analysis

Karen Wolf[1,2], *Pierre Fernique*[1], *Hans-Peter Piepho*[2]

[1]Limagrain Europe, France
[2]University of Hohenheim, Germany

To efficiently select plant varieties, breeders rely on accurate estimations of their performance (i.e., estimated genotypic value (EGV)). The accuracy of EGVs is largely influenced by field effects. When prior knowledge about field effects is available, breeders can account for them in the experimental design and estimate them without confounding with the EGVs. However, full prior knowledge of field effects is seldom available in practice, leading to deficient experimental designs.

We studied the impact of deficient experimental designs and their analysis on the accuracy of EGVs. To do this, we simulated field trials with varying field effects and designs. For the analysis, we used either the model implied by the experimental design (i.e., design-based analysis) or model selection (i.e., data-driven analysis). Our simulations are based on the distributions of field effects and blocking structures that we obtained from the analysis of historical field trial data. In general, we observe lower EGV accuracy if blocking does not account for all existing field effects. Moreover, for design-based analysis, EGV accuracy is negatively affected when blocks are used despite the absence of corresponding field effects. However, even for data-driven analysis, where such block effects can be dropped from the model, we may still encounter reduced EGV accuracy. For the experimental design, we present blocking structures which avoid negative impact on EGV accuracy despite lacking prior knowledge. For analysis, we recommend data-driven analysis when the data indicates that the blocking structure does not correspond to the observed field effects.

Designing multi-environment reference populations in perennials: do’s and don’ts

Michaela Jung[1], *Hélène Muranty*[1,2], *Andrea Patocchi*[1]

[1] Agroscope, Waedenswil, Switzerland
[2] Univ Angers, Institut Agro, INRAE, IRHS, SFR QuaSaV, F-49000 Angers, France

The multi-environment reference population (REFPOP) concept in perennial species was developed to promote genetic research, particularly for dissecting genotype-by-environment interactions in quantitative traits. It aims to address climate-related challenges and support the breeding of resilient cultivars. A decade ago, successful international collaboration led to the creation of the first European apple REFPOP. This effort faced challenges, including germplasm propagation and exchange, orchard space requirements, and long-term maintenance. Despite these obstacles, coordinated phenotyping across multiple sites over multiple years enabled comprehensive analysis of genotype-by-environment interactions and genomic prediction for 30 quantitative traits related to phenology, yield, and fruit quality. Key traits were accurately predicted in breeding material using genomic prediction and numerous novel marker-trait associations were discovered. Here, we outline crucial design elements of the apple REFPOP that contributed to precise genomic prediction and genome-wide association studies, in particular the incorporation of various accessions of global origin and multiple families representing the diversity of European breeding programs. A wide range of phenotypes among the genotypes included proved essential. However, some limitations were noted, such as the partial exclusion of parental genotypes from the included families, which hindered analyses of parent selection using genomic prediction. Additionally, the population’s diversity, coupled with differences in the application of reduced pesticide or reduced irrigation management across sites, likely resulted in the weakened detection of genotype-by-environment-by-management interactions. Lessons learned from the apple REFPOP, both do’s and don’ts, will inform the design and implementation of the upcoming European pear REFPOP.

Power calculations for testing genetic correlation between pure- and mixture line effects in wheat breeding

Bjarne Nielsen, Peter Skov Kristensen, Pernille Sarup, Pernille Bjarup, Jihad Orabi

Cultivar mixtures has been increasingly popular and a third part of the wheat cropping area in Denmark is cultivar mixtures. Cultivar mixtures have several potential benefits such as increased yield and stability, and reduced severity of diseases. This project aims to investigate genetic and non-genetic benefits of cultivar mixtures in wheat. In the coming three years we will run trials that include both pure lines plots and mixture line plots, and all lines will be genotyped candidates from a breeding program for new wheat cultivars at Nordic Seed. The aim of this power calculation study is to suggest optimal experimental design for estimation of the genetic correlation between the pure line genetic effects and the mixture line genetic effects. Mixtures of three wheat lines are considered, and for about 25 new candidate lines there are many possible combinations between three wheat lines in the mixture plots. To select the most informative mixture plots and to estimate the number of replicates needed for each line we use the predicted breeding values of traits from pure line trails of the new candidate lines and thereby increases the genetic variation between the selected lines for mixture plots. We assume a statistical model that simultaneously estimate genetic and non-genetic co-variances of pure- and mixture line effects. Our results suggest a selection strategy of lines and number of replicas that increases the power of testing the significance of the additive genetic co-variances between pure line and mixture line effects. The strategy will contribute to the development of more robust models that predict more accurate breeding values for wheat lines used for cultivar mixtures crops.

Session 3: 13:50-15:30

Genetic and genomic modelling of plant breeding data

Chair: Tally Wright – NIAB, England

A genomic selection reality check from industry: what moves the needle in genetic gain

Pascal Schopp

IS

KWS Group, Germany

The introduction of genomic selection (GS) in both animal and plant breeding programs has revolutionized breeding methodology and significantly increased rates of genetic gain. GS enables breeders to enhance conventional breeding schemes in multiple ways. In this presentation, I will share key lessons learned from the effective implementation of GS in commercial plant breeding programs, using the corn breeding program at KWS as a primary example, supplemented by insights from other crops within KWS' product portfolio. Beginning with the initial assumptions made during the early stages of GS adoption, I will explore how our understanding of its impact, opportunities, and associated risks has evolved over time. Which early expectations have been fulfilled, which have not, and what new insights have emerged through empirical learning in the context of large-scale industrial application? Specific examples will include strategies for scaling up breeding volumes and infrastructure, reducing cycle time, differentiating between prediction for population improvement and commercial performance, adapting phenotypic testing to support modern breeding schemes, improving selection accuracy under increasingly challenging environmental conditions, and managing genetic diversity in the era of prediction-based breeding. In addition, I will illustrate how the adoption of GS has transformed collaboration and workflows across various crop breeding programs and research initiatives at KWS. Finally, I will discuss the anticipated developments in genomic modeling and breeding methodology, and reflect on the evolving role of plant breeders and the significance of public-private research partnerships in this new landscape.

An efficient and powerful algorithm for detecting heterotic QTL

Guoliang Li, Renate H. Schmidt, Yusheng Zhao, Jochen C. Reif, Yong Jiang

To meet the growing global demand for food, increasing yields through heterosis in agriculture is crucial. A deep understanding of the genetic basis of heterosis has led to the development of a quantitative genetic framework that incorporates both dominance and epistatic effects. However, incorporating all pairwise epistatic interactions is computationally challenging due to the large sequencing depth and population sizes needed to uncover the genes behind complex traits. In this study, we developed hQTL-ODS, a one-dimensional scanning method that directly assesses the net contribution of each locus to heterosis. Simulations show that hQTL-ODS reduces computational time while offering higher power and lower false-positive rate. We applied this method to a population of 5,234 wheat hybrids with whole-genome sequenced profile, revealing key epistatic hubs that play a critical role in determining heterosis.

A composite hypothesis testing approach to detect pleiotropic genomic regions

Annaïg De Walsche[1,2], *Franck Gauthier*[2], *Nathalie Boissot*[3], *Alain Charcosset*[2], *Tristan Mary-Huard*[1,2]

[1]Mathématiques et Informatique Appliquées Paris-Saclay, AgroParisTech, INRAE, Université Paris-Saclay, 91120 Palaiseau, France
[2]Génétique Quantitative et Evolution - Le Moulon, INRAE, CNRS, AgroParisTech, Université Paris-Saclay, Gif-sur-Yvette, France
[3]Génétique et Amélioration des Fruits et Légumes, INRAE, Montfavet, France

Consider a study where one aims at detecting pleiotropy regions, i.e., the possible joint effect of a genomic region on (some or all of) K different traits. One seeks for a test procedure that rejects hypothesis H$_0$: “the region has no joint effect”, when at least two out of the K hypotheses H$_0$ “the region has no effect on trait k” are false, with $k = 1, ..., K$. This corresponds to a specific composite hypothesis that cannot be tested through conventional testing procedures. However, Composite Hypothesis Testing (CHT) can be performed by combining the p-values obtained for each of the K traits into a single summary statistics. This strategy has been successfully applied in human genetics to perform mediation analysis, to detect genomic regions associated with multiple diseases, or in the context of integrative genomics to analyse multiple omics traits. However most CHT procedures do not handle the dependencies between p-value series that may exist whenever the initial studies were performed on a same panel of individuals (e.g., hybrids) - a common situation in plant genetics. We propose a new CHT approach that explicitly accounts for the dependence structure across p-value series. The approach comes with an efficient implementation that alleviates the memory burden of the inference step, extending its application to series of millions of p-values and up to $K = 20$ traits. The effectiveness of our approach will be illustrated through simulations and through an application to the detection of hotspot regions linked to resistance to multiple viruses in cucumber.

Approximating prediction error variances of genomic models using Monte Carlo sampling

Antero Heikkilä[1], *Ismo Strandén*[2], *Martin Lidauer*[2], *Klaus Norhausen*[3], *Sara Taskinen*[1]

[1]Department of Mathematics and Statistics, University of Jyväskylä, Jyväskylä, Finland
[2]Natural Resources Institute Finland (Luke), Jokioinen, Finland
[3]Department of Mathematics and Statistics, University of Helsinki, Helsinki, Finland

Genomic prediction models such as GBLUP incorporate genomic data to improve the accuracy of estimated breeding values (EBVs). As the number of genotypes increase, the exact computation of EBVs and their reliabilities becomes computationally infeasible. Although iterative solvers such as the preconditioned conjugate gradient (PCG) algorithm can estimate EBVs without matrix inversion, they do not directly provide prediction error variances (PEVs) or reliabilities.

Our current research focuses on approximating PEVs and reliabilities in single-trait and multiple-trait genomic models using Monte Carlo (MC) methods. These methods avoid matrix inversion by repeatedly sampling breeding values from their assumed distributions and computing EBVs using iterative solvers.

We evaluated four existing MC approximation methods, originally developed for animal breeding, and adapted them for plant breeding applications. The convergence of the methods was assessed using varying levels of heritability, population size, and MC sample size, by comparing approximated PEVs to PEV values computed by (exact) direct inversion. Correlations between the exact and the approximated PEVs varied across the MC methods. These differences are largely explained by differences in sampling variance. All MC methods converged toward the exact PEVs as the number of MC samples increased.

Our results indicate that MC-based approximation methods are applicable in plant breeding, although their computational cost can be too large in very large-scale plant breeding.

Why REML variance estimates have no biological meaning and how to solve it

Javier Fernández-González, Julio Isidro y Sánchez

Centro de Biotecnologia y Genómica de Plantas (CBGP, UPM-INIA) - Universidad Politécnica de Madrid (UPM) - Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Campus de Montegancedo-UPM, Pozuelo de Alarcón, Madrid, Spain

Estimating variance components, particularly additive variance, is key to understanding trait architecture, heritability and making data-driven breeding decisions. Additive standard deviation is a core component of the breeder’s equation, and its accurate estimation directly impacts selection efficiency. Traditionally, restricted maximum likelihood (REML) in linear mixed models has been a go-to method. However, REML estimates are just scaling factors without biological interpretation, and using them as variance components can be misleading. Similarly, Bayesian methods that ignore uncertainty in posterior estimates also result in biased variance assessment. In this talk we will: i) demonstrate why REML is unsuitable for estimating biologically meaningful variance components, ii) review existing alternatives in the literature and discuss the advantages and disadvantages of each one and iii) Introduce a new, still unpublished, set of analytical equations that provides fast, accurate estimates of variance components along with their uncertainty, without relying on computationally expensive sampling from BLUP distributions. These equations require only a vector of BLUPs and their prediction error variances. The result is a precise estimate of the expected variance (i.e., the variance component) and its standard error, offering a practical and efficient solution for breeders and quantitative geneticists.

Session 4: 16:00-17:40

Innovative models with unconventional data streams

Chair: Laurence Moreau – INRAE, France

Beyond mixed models: integrative models for agricultural G2P using unconventional data streams

Hao Cheng

IS

University of California Davis, USA

Mixed models are a cornerstone of quantitative genetics, yet integrating them with novel data streams can unlock richer genome-to-phenome insights. In this talk, I introduce integrative modeling frameworks that combine mixed-model foundations with unconventional data, including images, textual descriptions, molecular phenotypes, and encrypted datasets, to improve genome-to-phenome (G2P) prediction in agriculture. I'll showcase new modeling approaches through applications such as end-to-end genomic prediction of visual and textual traits, a tiered data-sharing platform enabling collaborative research without exchanging raw data, and methods for incorporating molecular phenotypes such as transcriptomic, metabolomic, and epigenomic profiles. Additionally, I'll discuss approaches for cross-species annotation transfer and bidirectional methodologies bridging insights across species. Together, these integrative frameworks provide more accurate, scalable, and biologically meaningful G2P predictions.

Leveraging multi-omics for predictive modelling in perennial crops: a case study in grapevine

Hannah Robinson[1], *Maximilian Schmidt*[1], *Carlos Robles-Zazueta*[1], *Timo Strack*[1], *Peter Crisp*[2], *Hans-Peter Piepho*[3], *Kai Voss-Fels*[1]

[1] Department of Plant Breeding, Hochschule Geisenheim University, Geisenheim, Germany
[2] School of Agriculture and Food Sustainability, Faculty of Science, The University of Queensland, St Lucia, Australia
[3] Biostatistics Unit, Institute of Crop Science, University of Hohenheim, Stuttgart, Germany

Perennial crops pose distinct challenges for genetic improvement due to their regenerative nature, which contributes to complex genotype-by-environment interactions (GEI), alongside long generation times and high heterozygosity. Integrating multi-omics data offers a promising pathway to overcome these challenges and accelerate genetic improvement, yet its application in perennials remains limited. Here, we present findings from a multi-kernel prediction study in 210 Riesling grapevine clones, integrating 11 years of phenotypic data with 1.5M genomic (SNP), epigenomic (indirect methylation profiling), and phenomic (hyperspectral) data.

Our modelling shows that accounting for collinearity between omics-derived relationship matrices and applying non-linear kernels improves variance partitioning and the contribution to the clonal genetic effect. A multi-kernel approach using orthogonalized matrices achieved the best predictive performance. Back-solving kernel effects revealed that epigenomic features had the largest estimated impact on trait performance, highlighting their potential value for improving selection and gaining deeper insight into trait architecture.

Emerging analyses are investigating additive omics-by-environment interactions and non-additive effects such as intra- and inter-omic epistasis. While their predictive utility remains under evaluation, these models offer the potential to clarify biological signals often conflated in additive frameworks.

This case study illustrates the value of multi-omics modelling to improve prediction-based applications in perennial crops. As multi-omics data becomes more accessible, implementing reproducible, scalable pipelines will be critical. We outline future directions, including simulation-based optimization to guide multi-omics integration into perennial breeding programs, supporting genetic gain and long-term adaptation and productivity in the face of climate change.

Genomic prediction at the root: from one plant to better populations

Aiyesa Leke Victor[1], *Wolfgang Link*[1], *Stefan Scholten*[2], *Tim Beissinger*[3]

[1]Plant Breeding Methodogy Department, University of Göttingen, Germany
[2]Crop Plant Genetics Department, University of Göttingen, Germany
[3]Heritable Agriculture, Mountain View, California, United States

Genomic prediction (GP) in plant breeding has traditionally focused on inbred lines, structured families, and their progeny. However, no studies to date have explored the potential of using data from field-grown, individual plants (IPs) from heterogeneous populations as a training unit. Here, we assembled a panel of approximately 1,000 IPs derived from 50 diverse European maize landrace populations. Over two years, we collected phenotypic data for 12 agronomic traits and generated 120,924 SNP markers through genotyping-by-sequencing.

Our analyses using IP-trained prediction models revealed impressive prediction abilities for flowering traits (0.83–0.85), yield-related traits (0.39–0.50), and other architectural traits (0.55–0.82). Comparative evaluations across six linear and two machine learning models consistently showed that machine learning approaches outperformed traditional GBLUP methods. Moreover, cross-validation under four realistic breeding scenarios demonstrated the robustness of the IP-trained models in predicting out-of-year and out-of-population performance. We further showed how field experiments can be designed to minimize IP phenotypic error, and the optimal number of IPs and SNP-markers needed for such an IP-based prediction study.

These findings showcase the untapped potential of using IPs (with unrestricted genetic diversity) for genomic selection in outcrossing species. By improving early-generation selection in genetically diverse populations, this approach can significantly accelerate breeding cycles and enhance precision in complex trait improvement.

Quantifying the genetic basis of mixing ability and investigating its genetic architecture: case studies in wheat varietal mixtures and wheat-pea intercrops

T. Flutre, M. Remérand, J. Salomon, J. Enjalbert

Université Paris-Saclay, INRAE, CNRS, AgroParisTech, GQE, Le Moulon, Gif-sur-Yvette, France

Under the constraints of climate change and biodiversity collapse, a major agroecological lever consists in decreasing the use of synthetic inputs and diversifying farmers' practices to reduce the risk of crop failure. To this end, farmers often opt for growing varietal and crop mixtures. However, breeding programs work mainly with monovarietal stands only, despite genotype-by-stand type interactions.

We propose here to jointly analyze both monovarietal and mixed stands to understand the sources of decorrelation of genotype performance and the shift in plant-plant interactions between stand types. Coupled to the use of incomplete designs, such a strategy provides a solution to the combinatorial explosion, and leveraging genomic relationship matrices further improves the estimation of genetic (co)variances and prediction of direct and social breeding values (poster from Salomon et al.).

With this framework, we analyzed two kinds of field trials mobilizing hundreds of wheat genotypes over several seasons, the first involving varietal mixtures in alternate ranks (poster from Remérand et al.), the second involving crop mixtures with pea testers (poster from Salomon et al.). Beyond the estimation of genetic (co)variances, genome-wide association studies can be done on various agronomic (e.g., grain yield) and functional (e.g., specific leaf area) traits, for which we will show promising preliminary results.

This leads us to the study of ecophysiological models with genotype-dependent parameters, to better understand the genetic basis of compensations and synergies occurring in plant mixtures.

When durum wheat grows together with alfalfa: the genetic dialogue of neighbouring plants

Nicolas Salas[1], *Mikhail Beloglasov*[1], *Clara Voiry*[1], *Christine Tollon-Cordet*[2], *Bernadette Julier*[3], *Bruno Bernazeau*[2], *Alain Baranger*[2], *Jacques David*[1]

[1]Institut Agro Montpellier, UMR Genetic Improvement and Adaptation of Mediterranean and Tropical Plants (AGAP)
[2]INRAE, UMR Genetic Improvement and Adaptation of Mediterranean and Tropical Plants (AGAP)
[3]INRAE, Unité de recherche Pluridisciplinaire Prairies et Plante Fouragères

Crop diversification through intercropping optimizes resource use and fosters beneficial plant-plant interactions. However, designing productive species mixtures requires understanding the genetic mechanisms underlying such interactions, which remains underexplored despite their potential to shape community outcomes.

We investigated genetic interactions using an intercropping model of *Triticum durum* (wheat) and *Medicago sativa* (alfalfa), representing a complementary cereal-legume system. Our individual-centered design involved 6,400 plants from 181 wheat genotypes and 106 alfalfa half-sib families. Each plant was phenotyped for multiple traits including NIR spectra within unique intra- and interspecific neighborhoods, allowing estimation of direct and indirect genetic effect variances within and between species, and their covariances.

We applied quantitative genetic models estimating direct genetic effects (DGEs)-due to an individual's genotype-and indirect genetic effects (IGEs)-due to neighboring genotypes (Bijma et al.,2007; Muir, 2005). Our approach models how wheat genotypes affect neighboring wheat and alfalfa plants, and vice versa. We combined experimental and simulation approaches to estimate these effects and assess estimation reliability.

Across functional traits (tiller number, plant height, leaf nitrogen), agronomic traits (biomass, yield components), and NIR spectrums, we demonstrated the importance of accounting for genetic interactions and successfully estimated them for co-breeding applications. We detected significant pseudo-design heritabilities for both DGEs and IGEs across multiple traits. Simulations confirmed estimation accuracy and potential use in selection programs.

By formalizing multispecies genetic interactions, our work bridges quantitative genetics and agroecology, offering novel perspectives for breeding programs and illuminating facilitation mechanisms in crop mixtures.

Poster session 1 & Welcome reception: 17:40-19:40

Wednesday concludes with poster session 1 and the welcome reception. Refer to page 78 for abstracts.

Thursday 18th September

Session 5: 9:00-10:40

Genotype performance, stability and resilience

Chair: Katharine Preedy – BiOSS, Scotland

G x infectious E: breeding for reduced pathogen transmission

Andrea Doeschl-Wilson

IS

The Roslin Institute, United Kingdom

Infectious pathogens pose significant threats to the health of plants, animals, and humans. Reducing disease transmission is critical for enhancing disease resilience at the population level. While substantial progress has been made in understanding genetic resistance or tolerance to disease, much less is known about the genetic factors that influence pathogen transmission. In this talk, I will present both established and novel methods that integrate quantitative genetics with epidemiological models to estimate host genetic effects on disease transmission. I will illustrate these approaches using data from transmission experiments and field studies across multiple livestock species, and conclude by exploring their potential applications for managing infectious diseases in plants.

A hurdle model for ordinal scoring data with an underlying percentage scale

Emilia Koch, Hans-Peter Piepho, Jens Hartung, Benjamin Eppler

Ratings or scorings are often used, for example in ecology to score soil erosion, or in agricultural sciences for crop diseases, tackling crucial problems like yield loss effects. The resulting scoring data are strictly ordinal and often have an underlying percentage scale. Which model to use for this type of data is not straightforward to decide. Ordinal scores do not fit the assumptions required for analysis of variance. The threshold model, which was developed for ordinal data would be suitable, but it does not consider the underlying percentage scale. Hence, we here propose a model that fills this gap, a hurdle model for interval-censored percentage data. It models the data according to its nature: In its first part, it models presence or absence of a disease, and in the second part severity or abundance. Individually modeling presence and absence, allows to account for zero inflation. The second part borrows theory from the threshold model and the Johnson SB system of distributions that involves a transformation of the percentage scale to a normal distribution. The model result also reflects the two components. They individually describe the degree of disease infestation, and the degree of disease spread. This makes the result well-interpretable and enables insightful and concrete conclusions. To illustrate the model, we used mildew scorings from an on-farm trial in grapevines. We found the model highly suitable for this dataset, and superior to the threshold model.

How many environments do we really need for yield stability prediction in cereals?

Pablo Sandro[1], *Justin Blancon*[2], *Lucia Gutierrez*[3]

[1]University of Wisconsin - Madison, Madison, WI, USA
[2]INRAe, Clermont Ferrand, France
[3]Swedish Agricultural University (SLU), Alnarp, Sweden

Climate change poses a challenge for agriculture by increasing climatic variability and compromising crop yields, therefore, breeding for genotypes adapted to climate change is needed. One of the most important challenges for stability evaluations is that most stability indices require complete datasets and evaluations in large number of environments, while most breeding programs generate sparse multi-environment datasets. Methods such as Finlay-Wilkinson random regression can deal with sparse datasets and incorporate genomic data to leverage phenotypic information from related genotypes to predict yield stability. Our objective was to compare Finlay-Wilkinson random regression with genomic GBLUP models to predict yield stability, and to evaluate the impact of the number of environments and environmental variance on those predictions. We used three large datasets, one highly unbalanced dataset for oats, and two balanced datasets in barley and wheat. The Finlay-Wilkinson random regression model had higher predictive ability for grain yield for un-phenotyped genotypes (i.e., new genotypes) compared to the genomic GBLUP model in both balanced and unbalanced datasets. Because the Finlay-Wilkinson random regression model borrows information from relatives, the number of environments needed to have a high predictive ability was lower than using classic Finlay-Wilkinson models. On the other hand, pairs of contrasting environments, that created large among environmental variance, consistently yielded high predictive ability compared to classic Finlay-Wilkinson models indicating that a smaller number of contrasting environments could be used as training populations when random regression models are used.

The role of stability metrics in plant breeding programmes

Dominic Waters[1], *Christina Finegan*[2], *Chris Gaynor*[2], *Gregor Gorjanc*[1], *Christina Lehermeier*[2], *Brian Gardunia*[2], *Daniel Tolhurst*[1]

[1]The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, United Kingdom
[2]Bayer CropScience, St Louis, USA

Genotype by environment interaction (GEI) increases the complexity of selection decisions for both breeders and growers. With significant GEI, genotypes which combine high mean performance with stable ranking across environments are desirable. Most of the focus in research has been on developing metrics which capture differences in stability, with a view of ranking genotypes for variety release. In contrast, there has been very little research into how stability metrics could be utilised for parent selection in breeding programmes to simultaneously improve mean performance and stability. Breeders are keen to fortify their populations against variable environments and are already considering the use of indices which penalise genotypes with unstable performance. Before committing to selection decisions, it is important to review the quantitative genetic theory underlying GEI and differentiate between selection for improving a population versus releasing a variety to growers.

This talk will present an exploration of the impact of different strategies for improving mean performance and stability under different levels of GEI. We will demonstrate optimal crossing strategies for developing high performing and stable varieties and highlight key differences to strategies for improving the population. The results will provide a theoretical basis for the implementation of novel selection strategies in plant breeding programs and offer insight into the quantitative mechanics of selection in the presence of GEI. The tools developed in parallel to this research will assist breeding programs in effectively utilising GEI as a source of genetic variation from which to create, develop and release stable varieties.

Estimating genetic performance and variability on-farm using ranking data

Hugo Dorado[2], *Julian Ramirez-Villegas*[2,3], *Fred van Eeuwijk*[1], *Joost van Heerwaarden*[1,2]

[1]Mathematical and Statistics Methods, Wageningen University
[2]Plant Production Systems, Wageningen University
[3]Alliance of Bioversity–CIAT, Food, Environment, and Consumer Behavior, Cali, Colombia

Plant breeding research has chiefly relied on on-station experiments for the evaluation of varietal performance, but this has clear limitations in representing on-farm growing conditions and farmers' preferences, potentially leading to the misspecification of breeding targets. Recent work has demonstrated the potential of using simple, on-farm variety trials in combination with ranking data to evaluate (multi)-trait performance efficiently across multiple environments. To date, there has been a lack of methods for quantifying differences in genetic performance and trial quality from such trials (Individual trials/networks). We present a simple method for estimating genotypic means, genetic variance, and heritability using ranking data. We adapt a two-stage approach in which the estimated genotypic values and their standard errors are calculated using a model for ranking data analysis, then these estimates are used within a mixed model effect to estimate genetic standard deviation and heritability in the second stage. The performance and robustness of this approach is assessed using both simulated and observed data, including information on maize, cassava, and groundnut. The results demonstrate that ranking data can accurately estimate standard genetic parameters commonly used in plant breeding.

Session 6: 11:10-12:50

Modelling and leveraging genotype by environment interaction

Chair: Emi Tanaka – Australian National University, Australia

Exploiting genotype by environment interactions to breed for local adaptation and climate change resilience in winter rapeseed

Steven Penfield

IS

John Innes Centre, United Kingdom

Crop development can be divided into sequential stages, and local adaptation optimises the duration and the climate optima of each stage such that it matches that of the most frequent conditions at the desired growing location. In the modern era the availability of well-organised weather data can improve genomic selection for specific environments, but in order to decide what weather data to collect we need to understand which crop stages are weather sensitive and moreover, how this sensitivity relates to the prevailing variation in weather in target climate zones.

In European winter rapeseed the primary feature of local adaptation is the requirement for high freezing (‘winterkill’) tolerance in the east versus adaptation to warmer maritime winters in the west. In this presentation I will show how our desire to understand causes of yield variation in the maritime UK climate challenged our most basic assumptions of crop physiology and led to a new understanding of rapeseed growth stages and the role of winter chilling in arable crop yields. Further, via integration of stage-targeted weather data and public trial data I will show how we could identify varieties with increased resilience to variation in chilling and use these to discover genetic variation that can deliver adaptation to low chill environments.

Characterising genotypic stability and environmental responsiveness in Australian chickpea

Shanice Van Haeften[1], *Kristy Hobson*2, *Dominic Waters*[3], *Lee Hickey*[1], *Millicent Smith*[1,4], *Daniel Tolhurst*[3]

[1]Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Brisbane, Australia
[2]Department of Primary Industries and Regional Development, Tamworth, Australia
[3]The Roslin Institute, The University of Edinburgh, Easter Bush, United Kingdom
[4]School of Agriculture and Food Sustainability, The University of Queensland, Gatton, Australia

The increasing frequency of extreme weather events due to climate change poses significant challenges to global chickpea production, necessitating the development of more resilient varieties across diverse growing regions. Chickpea is ranked third in global production amongst pulses and is cultivated in over 50 countries. Despite its importance, chickpea production is highly variable across seasons and environments due to its sensitivity to diverse environmental conditions. Understanding how genotypes interact with the environment enables breeders to target breeding strategies for specific adaptation or broader stability across production regions.

This study builds on existing multi-environment trial analyses by integrating biologically meaningful environmental covariates, particularly those linked to heat. Three sets of environmental covariates were derived from weather data: i) phenology-aligned, ii) crude seasonal averages and iii) compound stress indices. Using historical yield data from Chickpea Breeding Australia (CBA) during 2018–2023, we demonstrate how our approach provides a more interpretable dissection of genotype by environment interaction compared to fully latent models by identifying environmental drivers of genotype re-rankings between environments, known as crossover interaction. This provides a platform to assess the genotypic stability and environmental responsiveness of breeding material developed by CBA, revealing genotypes which exhibit consistent performance across diverse environments or resilience to specific extreme weather conditions and stresses.

Overall, the findings highlight the complex nature of chickpea adaptation (exhibited through historical breeding data), and that utilising multiple biologically meaningful environmental covariates to inform selection decisions can be more effective than focusing on particular stress variables. The research provides valuable insights into breeding strategies for developing climate-resilient chickpea varieties, supporting sustainable production in a changing climate.

Multiple-covariate random regression for predicting yield across the Australian wheatbelt

Jip Ramakers[1], *Martin Boer*[1], *Jesse Hemerik*[1,2], *Daniela Bustos-Korts*[1,3], *Salvador Gezan*[4], *Vivi Arief*[5], *Javier Fernandez*[5], *Scott Chapman*[5], *Fred van Eeuwijk*[1]

[1]Biometris, Wageningen University & Research, Wageningen, the Netherlands
[2]Econometric Institute, Erasmus University, Rotterdam, the Netherlands
[3]Facultad de Ciencias Agrarias y Alimentarias, Universidad Austral de Chile, Valdivia, Chile
[4]VSN International, Hemel Hempstead, England, United Kingdom
[5]School of Agriculture and Food Sustainability, The University of Queensland, St Lucia, Australia

Environmental covariates (ECs) that capture G×E effects can be used to predict yield in new environments. An attractive model for this purpose is random regression (RR), which models random interactions between genotypes and ECs. These models, however, quickly run into fitting problems when the number of ECs > 1.

We have developed a RR model that allows fitting multiple ECs (including their higher-order polynomials) simultaneously to predict yield in novel environments. Using a factor-analytic structure on the RR coefficients (e.g., akin to Tolhurst et al., 2022), the genetic variance-covariance matrix - covering intercepts and slopes - can be readily expanded to accommodate >10 carefully selected ECs. Moreover, we can combine different RR structures for different environmental groups (e.g., zones) and borrow information between them.

Applied to 935 trials across the Australian wheatbelt, our multi-EC RR models achieved high median predictive accuracies (>0.6) in leave-one-year-out cross validation, 20–30% better than single-EC RR or standard compound-symmetry models. Combining different RR structures further boosted accuracies (4 %pts).

Our RR models are conceptually straightforward, computationally efficient, and easily implemented in ASReml-R, making them a practical tool for yield prediction when environmental data are available.

REML implementations of kernel-based multi-trait, multi-environment genomic prediction models

Killian Melsen[1], *Salvador Gezan*[2], *Fred van Eeuwijk*[1], *Carel Peeters*[1]

[1]Biometris, Wageningen University & Research, Wageningen, the Netherlands
[2]VSN International, Hemel Hempstead, England, United Kingdom

Plant breeding datasets spanning multiple years or locations are often incomplete in the sense that not all genotypes are evaluated in all environments. In this sparse testing setting, a small number of check varieties typically provide connectivity between environments. This lack of connectivity can complicate the estimation of genetic correlations between environments, often resulting in lack of model convergence, inaccurate estimates of genotype by environment interactions and thus inaccurate predictions of genetic values.

We show how weather station or open-source satellite data on environmental covariables can be used to accurately model genotype by environment interactions, resulting in better predictions of genetic values. We use linear and non-linear functions of these covariables, combined with environment specific or averaged genetic variances. Finally, we combine the use of environmental covariables with the unstructured modelling of multiple managements or traits, paving the way for easy integration of phenomics, enviromics, and genomics within the linear mixed model.

We implement all proposed models in standard restricted maximum likelihood based mixed model software. We also provide flexible, ready-to-use code that implements the proposed models for an arbitrary number of environments, managements and traits with no manual tuning required.

Tailoring AI and ML models for genotype-by-environment prediction leveraging environmental covariates: a European rye example

Wera Eckhoff[1,2], *Florence Parat*[3], *Gennady Bracho-Mujica*[2], *Clemens Flamm*[3], *Daniela Bustos-Korts*[4], *Hans-Peter Piepho*[1]

[1]Biostatistics Unit, Institute of Crop Science, University of Hohenheim, Stuttgart, Germany
[2]KWS SAAT SE & Co. KGaA, Einbeck, Germany
[3]Österreichische Agentur für Gesundheit und Ernährungssicherheit GmbH, Wien, Österreich
[4]Faculty of Agricultural Sciences, Universidad Austral de Chile, Valdivia, Chile

Accurate prediction of environment-specific genotype performance remains a major challenge in crop improvement and agricultural decision-making. ML and DNN-based predictions are often dominated by environmental main effects rather than genotype-by-environment interaction (GxE) effects which hinders their adoption in plant breeding. This study aimed to develop methodology to improve GxE predictive modeling in rye but also generally across crops by developing novel approaches to tailor ML and DNN models towards predicting environment-specific genotype differences and rankings rather than absolute performances. We introduce two methodologies: (i) target variable engineering based on linear mixed model (LMM) decompositions to isolate GxE effects, and (ii) a custom loss function implementation of the mean squared error of differences (Piepho, 1998) to optimize models directly for within-environment genotype difference prediction. The motivating dataset covers major rye growing regions worldwide and models were evaluated using a comprehensive cross-validation (CV) framework. Our approaches improved predictive abilities of ML/DNN models by +21 to +62 % compared to classical ML/DNN-based yield prediction. We increased predictive accuracy for environment-specific genotype rankings between +15.0% and +9.8% across different CV schemes over baseline genotypic main effects. Furthermore, leveraging historical weather records enabled prediction of genotype performances in future, untested years. This work demonstrates that tailored ML and DNN strategies can outperform classical methods for GxE prediction, offering practical value for plant breeders, growers, and variety testing authorities.

Social events & Conference dinner: 13:20-23:30

Thursday afternoon is reserved for the social programme. Refer to our website for details: `https://highlanderlab.github.io/EUCARPIA2025BiometricsPlantBreeding/social/`. The evening will include the conference dinner and cèilidh at Mansfield Traquair from 18:00.

Friday 19th September

Session 7: 9:00-10:40am

High-throughput phenotyping

Chair: Richard Mott – University College London, England

Spatio-temporal modelling of genotypic performance in high-throughput phenotyping

María Xosé Rodríguez-Álvarez[1], *Diana Marcela Pérez-Valencia*[2], *Martin Boer*[3], *Fred van Eeuwijk*[3]

IS

[1]CITMAga, Centro de Investigación e Tecnoloxía Matemática de Galicia and Universidade de Vigo, Spain
[2]Universidad de Deusto, Spain
[3]Biometris, Wageningen University and Research, The Netherlands

In agricultural experiments, phenotypic traits are often influenced by spatially structured non-genetic variation, such as soil heterogeneity or environmental gradients. Statistical spatial models, combined with appropriate experimental designs, have long been used to adjust for such variation and improve the estimation of genetic effects. With the advent of high-throughput phenotyping (HTP) platforms, phenotypic data can now be collected at high spatial and temporal resolution, offering valuable opportunities to study plant development and genotypic performance over time. However, the joint spatial and temporal structure of HTP data presents new challenges for statistical modelling, particularly in disentangling genetic trends from structured non-genetic variation in both space and time.

In this talk, we introduce a flexible modelling framework that extends spatial correction methods based on two-dimensional P-splines to the spatio-temporal setting. The approach uses three-dimensional P-splines within a mixed model framework to adjust for spatial and temporal variation in repeated measurements, while simultaneously modelling the temporal evolution of genetic effects via genotype-specific smooth curves.

Designed to be broadly applicable and computationally efficient, the framework can be used across a range of phenotyping platforms, from controlled environments to field trials. Examples from HTP experiments will illustrate the model's ability to estimate time-varying genotypic effects and to summarise and compare dynamic genotypic performance.

Bridging genomic prediction and crop growth modeling with System Identification for Nonlinear Dynamics (SINDy)

Corné Verburg, Alexander Heinlein, Fred van Eeuwijk, Neil Budko

The interaction between genetics and environment (G×E) is a fundamental challenge in plant breeding, complicating the prediction of crop phenotypes across diverse environments. Traditional genomic prediction models are static, relying on data from a single time point. However, advances in high-throughput technologies enable efficient collection of time-series data, presenting opportunities to develop models that better account for the temporal complexity of G×E interactions. Crop growth models (CGMs), such as WOFOST (WOrld FOod STudies) and APSIM (Agricultural Production Systems sIMulator), are based on first principles and incorporate years of plant physiology research. These models link environmental factors and crop traits. Unlike traditional statistical approaches using categorical environmental descriptors, CGMs employ environmental time-series data to simulate dynamic plant responses. However, CGMs face practical limitations in large-scale breeding programs due to numerous parameters and labor-intensive, high-quality data requirements for each genotype.

To address these limitations, we employ the Sparse Identification of Nonlinear Dynamics (SINDy) framework. SINDy is a model discovery method based on sparse regression that identifies symbolic expressions. It extracts dynamics from time-series data and describes crop growth curves using a small set of genotype-specific parameters. These parameters can be integrated into genomic prediction models, thereby merging mechanistic insights from high-throughput data with genomic prediction approaches. Using a dataset generated in silico via WOFOST, we demonstrate that SINDy successfully links the growth dynamics of leaf area index (LAI) and dry weight of storage organs to dynamic environmental conditions. This scalable, efficient approach shows great promise for modern plant breeding.

Deploying computer vision and bivariate genomic prediction to improve blackleg disease resistance in canola breeding

Timothy Thavarajah, *James Walter*, *Julian Taylor*

Brassica napus L. (canola/oilseed rape) is the world's second-largest oilseed crop, valued for its cooking oils, biofuels, and animal feed. However, blackleg disease (*Lepstophaeria maculans*) remains a persistent biotic stressor worldwide, causing 10-15% yield loss in Australia annually. Breeding for blackleg resistance relies on manual plant counting at emergence and physiological maturity in blackleg-infested nurseries to assess genotype survivability (derived from ratio of plant counts). This labour-intensive phenotyping approach limits genetic gain. Furthermore, selection for blackleg resistance is often performed through a univariate analysis of survivability, neglecting correlation between plant counts and disregarding their spatial and genetic variability.

This study demonstrates three key innovations: (i) high-throughput plant counting using computer vision; (ii) bivariate genomic selection (GS) using the repeated plant counts with a two-dimensional invariant multivariate autoregressive (2DIMVAR1) residual structure; (iii) a novel selection measure derived from genetic regression between emergence and maturity counts.

In collaboration with Australian Grain Technologies, we analysed multi-environment trials data at 4 locations across 3 years, along with over 9,000 genomic markers. Computer vision predicted emergence counts with high accuracy (R^2 = 0.95), offering a high-throughput, low-cost alternative to manual counting. Moreover, bivariate GS with flexible 2DIMVAR1 residual model more effectively modelled genetic and spatio-temporal covariance structures. Subsequently, breeders can screen for blackleg resistance within breeding population using their responsiveness to blackleg disease from implicit regression between genetic effects. These findings demonstrated the potential of integrating computer vision and bivariate GS to enhance phenotyping efficiency and selection accuracy for blackleg resistance in canola breeding programmes.

Integration of new RFID-based pollinator parameters into plant phenotyping and breeding selection indexes under climate change pressure

Ángela S. Prudencio[1], *Jose Ángel Martín-Rodríguez*[1], *Juan José Guerrero*[1], *Maialen Ormazabal*[1], *Cristina Martínez-Andújar*[1], *Purificación A. Martínez-Melgarejo*[1], *Antonio R. Jiménez*[2], *Francisco Pérez-Alfocea*[1]

[1]Department of Plant Nutrition, CEBAS-CSIC, Campus Universitario de Espinardo 25, 30100 Murcia, Spain
[2]Centre for Automation and Robotics (UPM-CSIC), Carr. Campo Real, 28500 Arganda del Rey, Madrid, Spain

Water scarcity combined with rising temperatures, both exacerbated by climate change, mainly disrupt the delicate balance in ecosystems, affecting crop physiology during flowering and ecosystem services such as plant-pollinator interactions. Today's digital phenotyping platforms are focused on genetic traits expressed in leaves. Indeed, the metabolism of flowers, and thus the quantity and quality of floral rewards (pollen and nectar), which serves as food for pollinators, change in response to the environment. We hypothesize that tracking and ranking pollinators' preferences for flowers under environmental pressure could be used as a marker of plant quality for agricultural breeding to increase crop stress tolerance. In this project, a phenotyping platform that integrates 336 Radio Frequency Identification (RFID) mobile antennas surrounding floral trusses and 48 data collectors was used to monitor tagged Bombus terrestris activity in six independent greenhouse modules. More than one million RFID detections were received and, after applying several data filtering criteria, were distilled into 14,288 confirmed pollination visits. In addition to Visit Number (VN), the RFID-tracking system let quantify each single bumblebee x plant interactions, rendering accumulated Total Visit Time (TVT), average Visit Time (VT) and number of Distinct pollinators -Tags- (DT) that visit each GxE combination. These four parameters were used to select candidate genotypes among a 30-lines tomato NIL (Near Isogenic Line) mapping population. Targeting floral traits in this way by breeding programs can help increase crop tolerance to abiotic stresses and would ensure food security through resilient crops and pollinator services.

Optimizing GxE prediction in phenomic selection using chemometric decomposition of NIRS data

Clément Bienvenu, Hugues de Verdal, David Pot, Vincent Segura

Phenomic prediction (PP) is a promising breeding tool which uses near infrared spectroscopy (NIRS) of plant tissues to predict breeding values. It uses the same models as genomic prediction (GP), but data collection is higher throughput and less expensive. Moreover, as a phenotypic measurement, NIRS can capture environmental signal that can be used to predict genotype by environment interactions (GxE). This feature could be an advantage for predicting genotypes' performances in multi environment trials.

It is possible to decompose spectra into genetic, environmental and GxE parts through modelling. However, GxE predictions with PP have only been done without such decomposition or using only the genetic part of the spectra. In this study, we investigated the relevance of properly decomposing the GxE parts of the spectra to predict genotypes' performances in multi environment trials. Spectra decompositions were done using linear mixed models or by adapting the LiMMPCA method from the chemometrics literature. Analyses were carried out on a rice diversity panel (188 genotypes) grown in four environments with four cross validation scenarios (sparse testing, leave genotypes out, leave one environment out, and leave one environment and genotypes out).

Results show a slight advantage of properly decomposing spectra for traits with low heritability in sparse testing and leave genotypes out scenarios. These encouraging preliminary results are currently being extended to other datasets to generalize these findings.

Session 8: 11:10-12:50

Artificial intelligence and machine learning in practice

Chair: Shanice Van Haeften – University of Queensland, Australia

Integrating artificial intelligence and machine learning capability into Australian grains agriculture

Julian Taylor

IS

University of Adelaide, Australia

Machine learning and artificial intelligence (ML/AI) are experiencing exponential growth globally across all industry sectors. Much of this is due to increasing computational power, step changes in model architectures and the synthesis of large and complex data sets. This growth is also matched by an increasing capacity of researchers entering the field at the technical and application level. It has spawned the postulation that ML/AI will solve all our problems, regardless of the industry but in most instances there needs to be a strong understanding of the physical domain where the problem was generated so targeted ML/AI solutions can be aligned and developed. In this talk, I will introduce the Analytics for the Australian Grains Industry (AAGI) project, and national industry funded initiative, where we have specifically included the Australian Institute for Machine Learning (AIML) as a key player to help integrate ML/AI technologies into Australian grains agriculture. The talk will focus on the complexities of ML/AI alignment in the grains industry as well as with other multi-disciplinary analytics that reside in the AAGI project. I will also showcase some examples of gains we have made using ML/AI technologies and touch on some AAGI research proposals and collaborations where we are attempting to drive future analytics innovation. I will conclude with some thought-provoking reflections on the evolving future of ML/AI and computing in agriculture, and the transformative possibilities that still lie ahead.

Exploring random forest-based genomic prediction for optimised selection of reference barley varieties for inclusion in DUS distinctiveness evaluation

Tally Wright[1], *Vanessa McMillan*[1], *Tansy Chia*[1], *Richard Horsnell*[1], *Malcolm Macaulay*[2], *Luke Ramsay*[2], *Paul Shaw*[2], *James Cockram*[1], *Margaret Wallace*[1]

[1] Niab, Park Farm, Impington, Cambridge, United Kingdom
[2] The James Hutton Institute, Dundee, United Kingdom

Shortening the time taken to release new crop varieties could increase the productivity and responsiveness of the breeding and farming sectors. To be awarded Plant Breeders' Rights a candidate variety must be shown to be morphologically distinct based on Distinctness, Uniformity and Stability (DUS) criteria from the Reference Collection – which for barley consists of thousands of genotypes. Currently this process takes a maximum of three years, using side-by-side comparisons of DUS phenotypes between the candidate and subsets of the Reference Collection. As candidates lack phenotype data in year-1 trials, subsets of the Reference Collections cannot be immediately identified for comparison. Genomic prediction could be used to shorten the testing cycle by selecting a prediction-informed subset in DUS year-1 trials. Here, DUS characteristic and genomic data were collected for 70% of the barley Technical Reference Collection, resulting in a representative training population consisting of 1,171 barley varieties, 28 DUS characteristics and 36,736 molecular markers. This dataset was then used to undertake Random Forest supervised machine learning to predict the DUS characteristics. The approach performed well compared to other prediction methods, and accuracy was increased via tuning hyperparameters and weighting major genetic loci. However, prediction accuracy was mostly dependent on trait heritability. Using evaluations with historical data, there was good crossover between the distinctiveness comparison sets selected via predicted or observed phenotypes, indicating the testing cycle could be reduced to a maximum of two years. Our findings highlight the potential of biomolecular technology to improve testing speed and efficiency in distinctiveness assessment.

Opportunities and limitations of deep learning-based multi-environment genomic prediction

Carles Quesada-Traver[1], *Steven Yates*[1], *Michaela Jung*[2], *Bruno Studer*[1]

[1]Molecular Plant Breeding, Institute of Agricultural Sciences, ETH Zurich, Universitätstrasse 2, 8092 Zurich, Switzerland
[2]Agroscope, Müller-Thurgau-Strasse 29, 8820 Wädenswil, Switzerland

Breeding resilient crop varieties demands tools that successfully account for genotype-by-environment (G×E) interactions. Multi-environment genomic prediction methods address this by integrating genomic and environmental data to estimate breeding values of genotypes in specific environments. Previously, we presented an interpretable multi-stream deep learning approach that successfully captured G×E using the apple REFPOP data. In this study, we aimed at testing the strengths and limitations of our modelling strategy alongside classical approaches using a different, larger dataset. For this purpose, we employed the Genomes-to-Fields initiative maize (*Zea mays* L.) dataset from 2014 to 2022. This dataset encompasses more than 70,000 phenotypic yield and flowering-related measurements in around 225 environments for more than 4,000 maize hybrids. Given the varying complexity and nature of G×E interactions across traits, we employed both grid search and Bayesian optimization to identify optimal hyperparameter (HP) combinations (such as network architecture, batch size and learning rate) tailored for each trait. These optimal HP combinations resulted in overall predictive abilities in the validation set ranging from 0.5 to 0.8 across traits. Furthermore, we present a new data-splitting strategy that enables validation of model performance across different real-life prediction scenarios. Our results demonstrate the suitability and flexibility of deep learning modelling to leverage vast amounts of data when performing multi-environment genomic prediction. We believe that in the long run, deep learning approaches can accelerate the development of cultivars adapted to future environments and promote sustainable agriculture through improved variety selection and more efficient resource use.

Image-based yield prediction for tall fescue using random forests and convolutional neural networks

Sarah Ghysels[1], *Bernard De Baets*, *Dirk Reheul*[1], *Steven Maenhout*[1]

[1]Department of Plants and Crops, Faculty of Bioscience Engineering, Ghent University, Ghent, Belgium
[2]Department of Data Analysis and Mathematical Modelling, Faculty of Bioscience Engineering, Ghent University, Ghent, Belgium

In the early stages of selection, many plant breeding programmes still rely on visual evaluations of traits by experienced breeders. While this approach has proven to be effective, it requires considerable time, labour and expertise. Moreover, its subjective nature makes it difficult to reproduce and compare evaluations. The field of automated high-throughput phenotyping aims to resolve these issues. A widely adopted strategy uses drone images processed by machine learning algorithms to characterise phenotypes. This approach was used in the present study to assess the dry matter yield of tall fescue and its accuracy was compared to that of the breeder's evaluations, using field measurements as ground truth. RGB images of tall fescue individuals were processed by two types of predictive models: a random forest and convolutional neural network. In addition to computing dry matter yield, the two methods were applied to identify the top 10% highest-yielding plants and predict the breeder's score. The convolutional neural network outperformed the random forest method and exceeded the predictive power of the breeder's eye. It predicted dry matter yield with an R^2 of 0.62, which surpassed the accuracy of the breeder's score by 8 percentage points. Additionally, the algorithm demonstrated strong performance in identifying top-performing plants and estimating the breeder's score, achieving balanced accuracies of 0.81 and 0.74, respectively. These findings indicate that the tested automated phenotyping approach could not only offer improvements in cost, time efficiency and objectivity, but also enhance selection accuracy.

Evolutionary-scale protein language models effectively uncover beneficial variants in a sorghum bicolor diversity panel

Natasha H. Johansen[1], *Janek Sven-Ole Sendowski*[2], *Eleni Nikolaidou*[2], *Savvas Chatzivasileiou*[2], *Baoxing Song*[3,4], *Shuai Wang*[3,4], *Andrew Olson*[5], *Doreen Ware*[5], *Thomas Bataillon*[2], *Guillaume P. Ramstein*[1]

[1]Center for Quantitative Genetics and Genomics, Aarhus University
[2]Bioinformatics Research Centre, Aarhus University
[3]Peking University Institute of Advanced Agricultural Sciences
[4]Shandong Laboratory of Advanced Agriculture Sciences in Weifang
[5]Plant Biology, Cold Spring Harbor Laboratory, Cold Spring Harbor

Quantitative genetic approaches including Genomic Prediction (GP) and Genome-Wide Association Studies (GWAS) are commonly used to identify favourable genetic variation. However, these methods suffer from low resolution due to linkage disequilibrium. Recently, deep learning methods, including Protein Language Models (PLM), have emerged as powerful tools for detecting phylogenetic residue conservation (PRC) across evolutionary time scales, possibly enabling more accurate variant effect predictions.

To identify potentially beneficial variants in the Sorghum Association Panel (SAP), we utilized the pretrained protein language model (ESM-v2) to predict PRC for nonsynonymous mutations within the population. The distribution of fitness effects (DFE) for these mutations was inferred using unfolded site frequency spectra (uSFS) to assess whether the DFE distribution was dependent on predicted PRC. Furthermore, the load of putatively beneficial variants, as identified from predicted PRC, was inferred for all accessions (n = 387) in SAP, and the associations between phenotypic performance and the genome-wide load of putatively beneficial variants were evaluated for multiple important agronomic traits to assess whether PRC can aid in detecting impactful variants for field traits. Results: We validated that ESM-v2 could efficiently detect mutations associated with fitness-enhancing effects in SAP. Furthermore, multiple significant associations between phenotypic performance and the load of putatively beneficial variants were detected for important agronomic traits, suggesting that PLMs may be effective tools for identifying non-neutral variants.

Our findings show that many quality traits were negatively associated with the load of putatively beneficial variants, indicating that artificial and natural selection may diverge.

Session 9: 13:50-15:30

Managing genetic diversity and maximising long-term genetic gains

Chair: Sanjeev Kumar Sharma – The James Hutton Institute, Scotland

Effective and efficient utilization of gene bank resources using quantitative genetics

Keith Gardner, *Sarah Hearne*, *The participants of the CGIAR Allele Mining project* **IS**

International Maize and Wheat Improvement Center (CIMMYT), Texcoco, Mexico

For over 60 years, CIMMYT maize and wheat breeding programs have been providing stable yielding, disease-resistant and climate-adapted varieties for resource-poor farmers around the globe. Novel genetic diversity for these programs is readily available from CIMMYT's germplasm bank, which hosts the largest collection of maize and wheat in the world: 28,000 accessions of maize and over 120,000 accessions of wheat landraces and wild relatives. I discuss how approaches to discovery and deployment of useful diversity from this collection have evolved over time, from ad hoc breeding efforts to systematic genetic screening, and from exploitation of major gene variants to use of quantitative variation. Specifically, I examine how quantitative genetic approaches have become increasingly important in both discovery and deployment and how we are currently using such approaches in ongoing 'allele mining' efforts.

A rapid cycling genomic selection experiment in maize landraces

Carolina Rivera-Poulsen[1], *Clara Polzer*[1], *Michelle Teran-Pineda*[1], *Armin C. Hölker*[2], *Thomas Presterl*[2], *Sofia Pereira da Silva*[2], *Manfred Mayer*[1], *Hans-Jürgen Auinger*[1], *Milena Ouzunova*[2], *Albrecht E. Melchinger*[1,3], *Chris-Carolin Schön*[1]

[1] Plant Breeding, TUM School of Life Sciences, Technical University of Munich, Freising, Germany
[2] Maize Breeding, KWS SAAT SE & Co. KGaA, Einbeck, Germany
[3] Institute of Plant Breeding, Seed Science and Population Genetics, University of Hohenheim, Stuttgart, Germany

Landrace material can serve as a source of novel genetic diversity for maize breeding programs. This requires closing the yield performance gap between landraces and elite germplasm, whether the target trait is yield or the enhancement of other quantitative traits through broader genetic diversity. In this study, we investigated the potential of a selected maize flint landrace adapted to Central European growing conditions to establish a pre-breeding program utilizing rapid cycling genomic selection.

The experiment comprises three cycles of recurrent genomic selection for testcross performance. As a training population for all selection cycles, data from a DH library derived from the flint landraces Petkuser Ferdinand Rot and Kemater Landmais Gelb was used (N = 419). Selection was based on a selection criterion comprising multi-trait genomic breeding values for total dry matter yield (TDMY), plant height (PH) and days to silking (FF). DH populations from all cycles were evaluated in multi-environment field trials, together with seven commercial hybrids as checks.

Over the three cycles of selection a significant increase in TDMY was observed. With a total increase of 10%. Thus, the performance gap to the commercial check hybrids changed from 23% in the initial population to 16% in the third cycle. Despite being under stabilizing selection, PH and FF also changed. Prediction accuracy for traits under selection was high in the initial cycle and subsequently decreased in the next cycle. This study highlights both the potential and challenges of applying rapid cycling genomic selection to maize landraces in a pre-breeding context.

An algorithm for mate selection by haplotype stacking

Kira Villiers[1], *Kai Voss-Fels*[2] *Ben J. Hayes*[1]

[1]The University of Queensland, St Lucia, Australia
[2]Hochschule Geisenheim University, Geisenheim, Germany

Genomic selection has proven to be a technique to increase rates of gain in breeding programs across varied agricultural species. Beyond simply using genetic information to cross the plants with the highest predicted trait values, the information can be used to design crossing plans that control rate of inbreeding (optimal cross selection, OCS (Kinghorn 2011)), or whose expected progeny distributions have high-scoring tails (usefulness criterion, UC (Schnell and Utz 1976) and its descendants) in order to increase rates of gain in the short or long term. Here we investigate whether crossing strategies inspired by gene stacking could further increase rates of gain. One strategy that shows promise in simulation is an evolutionary computing algorithm that designs matings between selected parents to bring together favourable haplotypes across the genome, increasing the length of "runs" of favourable haplotypes in the next generation. Emergent properties of this strategy include picking matings with lower coancestry and higher within-family variance, compared to random mating. Whether the parent population is selected for diversity or not, the use of this strategy increases genetic gain and reduces loss of genetic diversity in the simulated population. This proposed algorithm for designing a mating plan could therefore be a useful tool in the design of plant breeding programs.

The maximisation of the genetic gains of the polyclonal selection in grapevine varieties using integer programming

Sónia Surgy[1], *Jorge Cadima*[2], *Elsa Gonçalves*[1]

[1]LEAF—Linking Landscape, Environment, Agriculture and Food—Research Center, Associated Laboratory TERRA, Instituto Superior de Agronomia, Universidade de Lisboa, Lisboa, Portugal
[2]Instituto Superior de Agronomia, Universidade de Lisboa, Lisboa, Portugal

Polyclonal selection is a selection technique that was developed in the context of ancient grapevine varieties. It employs the predictors of the genetic effects obtained from the fitting of a mixed model. The term can be defined as the selection of a group of different genotypes which, when considered as a whole, meet certain criteria for desirable quantitative traits. The emphasis is placed on the collective behaviour of the group as a whole rather than on the individual behaviour of each genotype within the group. The identification of the superior group of genotypes may be regarded as an optimisation problem, with the objective being to maximise the predicted genetic gain of the group as a whole in accordance with the selection criteria adopted, considering several important quantitative traits. The objective of this study is to illustrate the implementation of Integer Programming in the context of polyclonal selection, and to present an algorithm that seeks to optimise the genetic gains of selection in accordance with various selection criteria. A package using the R programming environment is presented to support this work, and specific cases are tested with real data from selection field trials of ancient grapevine varieties. The methodology ensures that selected groups with high genetic gains in the target traits are obtained, while losses in other equally important traits are avoided. This procedure could be extended to plant breeding programmes for other species whenever the objective is to select a group of genotypes.

A novel framework to control genetic diversity for optimal genomic mating

Seifelden M. Metwally, Javier Fernández-González, Julio Isidro y Sánchez

Centro de Biotecnología y Genómica de Plantas (CBGP, UPM-INIA), Universidad Politécnica de Madrid (UPM) - Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Campus de Montegancedo-UPM, 28223-Pozuelo de Alarcón, (Madrid), Spain

The optimization of mating plans, or optimal genomic mating (OGM), is a well-established breeding strategy that balances genetic gain with population diversity, thus ensuring long-term breeding success. The aim is to find mating plans that optimally balance predicted family performance with two sources of variability in breeding populations: within- and between-family variance, under realistic practical constraints. Over the past 15 years, OGM methodologies have advanced significantly. Here, we introduce a novel method employing a standardized and intuitive metric (selection intensity) to control genetic diversity. This enhances the interpretability and comparability of different OGM strategies. We rigorously evaluated it through stochastic simulations across three breeding systems (line, clonal, and hybrid), covering diverse trait architectures, genetic effects, parental pool sizes, and generation intervals. Our proposed framework consistently and efficiently converted genetic diversity into realized genetic gains, promoting sustainable genomic selection programs for future genetic improvement.

Poster session 2 & Farewell reception: 15:30-17:00

Friday concludes with poster session 2 and the farewell reception. Refer to page 114 for abstracts.

Poster session 1

1. Disentangling the genetic and environmental factors influencing GxE for barley yield

Stephanie Brunner[1], *Zachary Aldiss*[1], *Samir Alahmad*[1], *Hans-Peter Piepho*[2], *Silvina Baraibar*[3], *Dini Ganesalingam*[3], *David Moody*[3], *Lee Hickey*[1], *Kai Voss-Fels*[4], *Hannah Robinson*[4]

[1]Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, St Lucia, QLD, Australia
[2]Biostatistics Unit, Institute of Crop Science, University of Hohenheim, Stuttgart, Germany
[3]InterGrain Pty Ltd, Perth, WA 6163, Australia
[4]Department of Plant Breeding, Hochschule Geisenheim University, Geisenheim, Germany

Barley is cultivated across a wide range of growing environments, making genotype-by-environment interactions (GEIs) a key consideration for breeders. As climate change increasingly threatens crop productivity, understanding the factors contributing to GEIs in barley yield has become critical. This study uses a robust data set 668 genotypes evaluated at 21 field sites across Australia to investigate methods for disentangling the GEIs at the genotype, haplotype and environment level.

Haplotype based methodology was used to identify genomic regions contributing to GEIs across a range of environmental clusters. GBLUPs from a multi-environment trial analysis with a factor-analytic variance structure applied to the GEI term were used to back-solve marker effects specific to each environmental cluster. The markers were then partitioned into 1,758 haplo-blocks based on linkage disequilibrium. Haploblocks underlying GEI for yield were then identified across all barley chromosomes. These blocks were co-located with previously reported key phenology and root QTL to aid interpretation of the complex genotype-by-environment relationship. Blocks co-located with root QTL were predominately associated with scale-type GEI for yield, whereas those linked to phenology typically exhibited cross-over GEI.

To facilitate interpretation of environmental influences on yield GEIs, a synthetic co-variate approach was applied to pinpoint key environmental drivers of yield performance and GEI across different growth stages. The outcomes of this study offer improved insight into the nature of GEI affecting barley yield and present a framework for application across species to better understand the complex interactions between genotypes and environments.

2. AI-driven analysis of big biological data to decode genotype-environment interaction

Huihui Li

Genotype, environment, and genotype-by-environment (G×E) interactions critically influence crop phenotypic formation. In this study, we developed and validated an automated machine learning framework integrating environmental and genomic data using a large-scale multi-environment hybrid maize dataset, which significantly enhanced the accuracy and efficiency of genetic analysis and genomic prediction. By employing developmental stage-aligned, dimensionality-reduced environmental parameters (RD_EPs), we established linear relationships between RD_EPs and traits to quantify environmental effects on phenotypes. Genome-wide association analysis identified 539 phenotypic plasticity-associated markers (PP-TAMs), 223 environmental stability markers (Main-TAMs), and 92 G×E interaction markers (G×E-TAMs), revealing distinct genetic bases underlying phenotypic plasticity and G×E interactions. The simultaneous incorporation of TAMs and RD_EPs into genomic prediction models improved prediction accuracy by 14.02-28.42% compared to genome-wide marker approaches. These results demonstrate the potential of environmental data in optimizing genetic analysis and genomic selection, providing a scalable research paradigm for developing climate-adaptive maize varieties. Furthermore, comparative model analyses revealed that the automated machine learning framework not only enhanced predictive performance but also substantially improved computational efficiency compared to conventional models, offering a novel and practical tool for rapid implementation of genome-wide selection breeding in crops.

3. From light to biomass: using dynamic photosynthesis data to improve biomass predictions

Junita Solin[1,2], *Tom Theeuwen*[2], *Martin Boer*[1], *Fred van Eeuwijk*[1]

[1]Biometris, Wageningen University & Research, The Netherlands
[2]Jan IngenHousz Institute, The Netherlands

Genetic variation of photosynthesis among and within crop species suggests opportunities for yield improvement and future breeding targets. When and how photosynthesis limits yield remain to be answered. Photosynthesis has been extensively studied under low-throughput, growth-chamber experiments, but only recently have imaging platforms allowed us to measure photosynthesis continuously, under dynamic environmental conditions. In contrast to agronomic traits such as plant height or projected leaf area, which typically follow well-defined patterns over the growing cycle and are directly related to yield, photosynthetic variables exhibit more complex dynamics and a less straightforward relationship with yield. This complexity presents challenges in longitudinal modelling and linking photosynthesis to yield.

Using the phenotyping platform in NPEC (Netherlands Plant Eco-phenotyping Centre), we could phenotype the photosynthetic activity of 394 Arabidopsis MAGIC lines every 15 minutes during the daytime (8.00-20.00) over 12 days. The climate chamber is set to a series of dynamic lighting and temperatures that mimic field conditions. We used p-splines (cyclic and non-cyclic) to estimate the temporal trajectories of photosynthesis variables at the population and genotype level. Our results show that photosynthesis variables changed diurnally according to the light conditions, with trends varying for different genotypes both within and across days. Following the smoothing, we extracted curve parameters and used them to predict biomass, the proxy for yield. Overall, this framework provides a useful approach to leverage time-series photosynthesis data to predict end-point traits, while also understanding the key aspects of photosynthesis that limit biomass or yield.

4. Elucidating the genetic architecture of heterosis in Central European wheat

Guoliang Li, Yong Jiang, Renate H. Schmidt, Jochen C. Reif

Despite decades of research and development, the commercial adoption of hybrid wheat remains limited, with hybrid cultivars occupying only a negligible share of the global wheat market. Advancing our understanding of the genetic basis of heterosis is therefore essential for designing more effective hybrid breeding strategies. To address this, we performed a systematic genome-wide mapping of heterosis and its underlying genetic components using a large-scale hybrid population generated from crosses among deeply sequenced inbred lines. We identified hundreds of heterotic QTL (hQTL) across the wheat genome, revealing that heterosis is governed by a highly polygenic architecture. Further dissection of these QTL showed that epistatic interactions, rather than simple dominance, are the predominant contributors to heterosis for grain yield, heading date, and plant height. Notably, we discovered an epistatic hub at the distal end of chromosome 4A coinciding with an alien introgression from emmer wheat. To further resolve the genetic basis of heterosis, we leveraged both triple testcross and immortalized F_2 designs, which enabled the detection of a major hQTL in the hybrid Piko × Hermann, again highlighting the central role of epistasis in heterosis expression. Epistatic interactions between hQTLs and the genetic background play a predominant role in shaping heterosis which provide a conceptual foundation for the development of heterosis-informed breeding strategies

5. Ideas and recommendations for optimal field experimental designs in artificial selection programs

Alexandre Colmant, Fabiano Pita, Giovanny Covarrubias-Pazaran

Experimental designs in classical and genome-based artificial selection programs are the cornerstone of high-quality phenotypic evaluations. Experimental designs run across and within locations have the final goal to ensure the maximum accuracy in the genetic evaluation by removing or reducing noise from genetic signal. Unfortunately, many programs only focus on specific decisions potentially sacrificing the accuracy of the overall evaluation. Approaching the resource-allocation problem holistically requires partitioning the breeding process into different decision points that can help to optimize each by separate but at the same time holistically. Here we present how to partition the process in key questions and give some recommendations on how to optimize them using well-suited objective functions and we back up our recommendations using stochastic genetic simulations finetuned for the different decision points. We found meaningful increases in accuracies and expected genetic gains when using the proposed objective functions. The increases in accuracy and genetic gain due to different objective functions varied depending on the complexity of the target population of environments (TPE) and genotypes (TPG) and the implications and hidden patterns are discussed. We hypothesized that our recommendations would help artificial selection programs to optimize their phenotypic evaluations and guide their decision-making.

6. Using phenomic selection to predict hybrid values of parental lines in nurseries - proof of concept on maize

Renaud Rincent, Junita Solin, Jérémy Labrosse, Adam Serghini, Marc Labadie, Alexis Comar, Laurence Moreau

Phenomic selection (PS) is a predictive approach similar to genomic selection but with molecular markers replaced by near-infrared (NIR) spectra or other high-throughput phenotypes. This low-cost approach has proven to be useful for many plant species. However, for hybrid species, direct application of PS to predict hybrid performances requires the production of the hybrids on which spectra will be acquired, meaning that PS cannot be applied at the early stages of the breeding programs. Our objective was to overcome this strong limit. For this, we recently proved, using an incomplete factorial design of maize, that spectra measured on the parental lines in a nursery (on leaves and grains) could be used to predict their hybrid values (GCA and SCA) in an independent multi-environment trial. We went a step further by using drone images of the nursery instead of NIR spectra, further increasing the throughput and decreasing the cost. The images were converted into secondary traits or summarized into latent variables using an autoencoder. The predictive abilities obtained with these approaches were competitive with those of genomic selection for complex traits, opening the way to the use of PS to predict complex traits in the nurseries. As shown in previous studies, PS seemed to be particularly useful to predict genetically distant material, a situation in which genomic selection failed.

7. Genomic and phenomic prediction performance for tree architecture and fruit quality traits in apple

Nuri Güvencli, *Hannah Robinson*, *Carlos Robles Zazueta*, *Kai Voss-Fels*

Hochschule Geisenheim University

As a perennial crop, apple (*Malus domestica* Borkhaus) breeding is notoriously slow. Predictive breeding can accelerate selection, but accuracy depends on trait genetic complexity, modeling strategies, omics data sources, and their preprocessing workflows. In this project, we systematically compare genomic, phenomic, and integrated multikernel approaches across multiple prediction contexts and cross-validation schemes for traits of varying complexity.

Four biparental populations segregating for columnar habit and anthocyanin expression were genotyped with a 48K SNP array. Hyperspectral measurements of leaves at key developmental stages and of fruit were conducted. Target traits included tree architecture (trunk length, slenderness), fruit quality (firmness, size, shape), appearance (skin and flesh colouration), and juice properties (acidity, soluble solids content). We implemented genomic prediction as a baseline, phenomic models with different preprocessing approaches for noise- and multicollinearity reduction, and multikernel frameworks integrating genomic and processed phenomic kernels. Models were evaluated under random cross-validation and leave-one-population-out scenarios to assess unbiased predictive performance across contexts.

Outlook: By benchmarking these approaches across trait categories, this study will inform on trait genetic architectures in apple, their impact on optimal model and omics choices, and optimal wavelength preprocessing strategies, guiding streamlined workflows for breeding pipelines.

8. Increasing nitrogen use efficiency of winter oilseed rape (*Brassica napus* L.) by improving genetics and cultivation system interaction

Daniel Valle Torres[1], *Sebastian Warnemünde*[1], *Nazanin Zamani-Noor*[2], *Milka Malenica*[3], *Christian Flachenecker*[4], *Amine Abbadi*[3], *Franz-Leopold Haupt*[5], *Jakob Streuber*[5], *Thomas Kreuter*[5], *Sven Weber*[6], *Benjamin Pommerrenig*[1], *Andreas Stahl*[1]

[1] Julius Kühn Institute (JKI), Federal Research Centre for Cultivated Plants, Institute for Resistance Research and Stress Tolerance, Quedlinburg, Germany
[2] Julius Kühn Institute (JKI), Federal Research Centre for Cultivated Plants, Institute for Plant Protection in Field Crops and Grassland, Braunschweig, Germany
[3] Norddeutsche Pflanzenzucht Hans-Georg Lembke KG (NPZ) Innovation GmbH, Holtsee, Germany
[4] Norddeutsche Pflanzenzucht Hans-Georg Lembke KG (NPZ), Holtsee, Germany
[5] SKW Stickstoffwerke Piesteritz GmbH, Cunnersdorf, Germany
[6] Justus Liebig University, Department of Plant Breeding, Giessen, Germany

For high crop productivity, agronomic management relies on substantial nitrogen (N) fertilization. However, plants do not take up the entire N applied, resulting in a limited N use efficiency (NUE). The factors genetic background, N fertilizer type and sowing technique influence NUE. Typically, their effects have been analyzed separately. By studying the interactions between genetics, environment, and agronomic management (G × E × M), we aim to identify optimal combinations for high NUE as well as underlying genetics. Breeding progress of NUE exploits genetic diversity by focusing on traits such as stay-green, plant architecture and primary yield components thousand seed weight, siliques per m^2, and seed to silique ratio. In three-year multi-environment trials in Germany, using fertilizers leading to either nitrate- or ammonium-based nutrition, testing cultivation varying on density and sowing techniques, we explore genetic variation in a collection of 300 F1-hybrids from an elite breeding program for winter oilseed rape. We investigate genetic and physiological determinants for high NUE and high yield, by integrating drone-based multispectral reflectance measurements, 360° RGB field images of developing plants and near-infrared spectroscopy data from seeds, into genomic prediction models. First year results identified 30 contrasting genotypes in terms of yield [3.3 - 4.7 t/ha], oil [47.4 - 51.0%] and protein content [16.6 - 19.5%]. Significantly, higher mean protein values were observed in ammonium (18.0 ± 0.1) compared to nitrate-based (17.7 ± 0.1) nutrition. Combining genomic and multispectral phenomic selection, we aim to predict genotypes with increased NUE while maintaining high yield.

9. Quality over quantity? The optimized allocation of quality samples of perennial ryegrass in Bavarian state cultivar trials

Anne-Katrin Gorn, *Jens Hartung*, *Stephan Hartmann*, *Hans-Peter Piepho*

In Germany, cultivars are tested for regional recommendations in federal state cultivar trials, taking the form of multi-environment trials (METs). Their primary objective is to identify cultivars that are best suited for regional production in agro-ecological zones. For perennial ryegrass, current selection decisions are predominantly based on yield. Incorporating additional quality characteristics could improve the selection process. In Bavarian state cultivar trials, additional quality parameters were taken for perennial ryegrass to describe cultivar quality. Due to financial constraints, the number of samples sent to the laboratory for analysis annually is limited. Consequently, single-plot samples were partly mixed across replicates and analysed as mixed samples. The decision which single-plot samples were mixed for mixed samples were made by intuition. The objective of the current study was finding an optimal distribution of single-plot and mixed samples for quality parameters in perennial ryegrass. The aim was to achieve the highest possible precision without surpassing the constraints of limited samples and workload. Data from METs of perennial ryegrass across three sites in Bavaria during the years 2017–2023 were analysed. The analysis comprises two main steps. First, variance components at each site across trial cycles were estimated. The second step involves the simulation of data representing alternative sampling designs. Data of alternative designs were analysed with variance components fixed to the estimates obtained in the first step. The precision measured as power and standard error of treatment differences was assessed for each alternative design. The results and their implications will be presented and discussed.

10. Breeding for maize varieties with reduced carbon footprints

K.R. Grant[1], *C.M. Richardson*[1], *T. Oliveira*[1], *L. Lara*[1], *C. Adams*[1], *M. Post*[1], *T. Byrne*[1], *P. Amer*[1], *W. Bourdoncle*[2], *R. Bouchon*[2], *L. Busswinkel*[2], *S. Larmer*[2], *B. Gardunia*[2]

[1] AbacusBio,
[2] Bayer Crop Science

In recent years, the agricultural sector has faced increasing pressure to reduce its carbon footprint. Genetic selection offers a permanent and cost-effective strategy for achieving this. This abstract presents how a commercial breeding program has adopted a novel approach to breeding maize varieties that are both high-yielding and environmentally sustainable, using a selection index.

Selection indexes are a useful tool that have been historically used by breeders and farmers to select the most profitable candidates. With some adaptions to selection index theory, it is also possible to select for sustainability. Sustainability selection indexes have already been developed and deployed by cattle breeding programs to reduce methane emissions. This is the first example of a selection index approach being used in a commercial plant breeding program.

Here, we will present how a selection index was developed to identify maize varieties with reduced carbon intensity emissions. The index incorporates multiple traits, including yield, nitrogen use efficiency, standability, and resistance to diseases. The resulting selection index can identify the most promising maize candidates for advancement. It can also be used to make predictions about genetic gain and the role of genetics in reducing the carbon footprint of maize.

The ultimate goal is to create an array of high-yielding, low-carbon maize varieties that can meet global food demands while minimising environmental impact. The use of a selection index in maize breeding represents a significant advancement in the pursuit of sustainable agriculture.

11. Leveraging disentangled representations to predict unobserved genotype-environment combinations in phenomic selection

Hugo Gangloff[1], *Do Than Dat Le*[1], *Renaud Rincent*[2], *Julie Aubert*[1], *Tristan Mary-Huard*[1,2]

[1]Université Paris-Saclay, AgroParisTech, INRAE, UMR MIA Paris-Saclay, 91120,Palaiseau, France
[2]Université Paris-Saclay, INRAE, CNRS, AgroParisTech, GQE - Le Moulon, 91190,Gif-sur-Yvette, France

Phenomic Selection aims at predicting genetic value of individuals (lines or hybrids) based on their Near-Infrared Spectroscopy (NIRS) spectrum. Because NIRS spectrum are collected within specific environments, they provide information about the individuals, environments and GxE interactions. This dual information may be challenging to exploit when focusing on inter-environment predictions, where the goal is to predict the performance of an individual in a target environment based on the phenotypes and NIRS from other environments.

The presentation will focus on a scenario involving multi-trial experimentation with a large panel of partially phenotype individuals. The goal is then to predict the performance of individuals in the environments where they have not been observed, an application case for which no dedicated methodological approach currently exists.

We proposed a Disentangling Conditional Variational Autoencoder (DC-VAE) approach, where “disentangling” refers to separating the information in the spectrum into components associated with the individual’s genotype and the environment, respectively. These components are then recombined in novel combinations corresponding to unobserved individual-environment associations to perform prediction. The proposed procedure makes minimal assumptions, requiring only that the experimental design is connected and that each individual to be predicted has data from at least one other environment. The DC-VAE approach will be demonstrated through an application to predict grain yield in maize using data from a four-environment trial.

12. Disentangling the genetic response to seasonal and environmental drivers

Katharine F. Preedy[1], *Brezo Mateos*[2], *Robert D. Hancock*[2], *Julie Graham*[2]

[1]Biomathematics and Statistics Scotland
[2]The James Hutton Institute

A changing climate is driving increased interest in the ability to breed for specific environmental conditions and an associated need for techniques to reliably identify genetic markers associated with key phenotypic responses. The degree environmental variation in field trials makes identification of cues expensive, time consuming and difficult. The size of populations required for conventional GWAS or QTL approaches is such that controlled environments are impractical experimental environments and only partially address the issues of seasonal drivers. Using dormancy induction in Raspberry as an example, we present a pipeline for analysing RNA-Seq data using integrating Weighted Gene Co-expression Network Analysis, Mixed Effects models, k-means clustering and constrained ordination followed by functional analysis to pinpoint the genetic basis for dormancy disruption in key raspberry cultivars. We further discussion results validation through metabolomic analysis and field scoring. The general approach is applicable to a variety of ‘omics data using widely available analytical tools.

13. Introducing the 2NP matrix in genomic prediction: a novel genomic matrix that merges the strengths of classical and machine learning methods in plant breeding.

Bright Enogieru Osatohanmwen[1], *Dr. Indalécio Cunha Vieira Júnior*[2], *Prof. Dr. A. Reza Sharif*[3], *Prof. Dr. Timothy Beissinger*[4]

[1]Department of Crop Science, Division of Plant Breeding Methodology, University of Goettingen, Germany
[2]KWS SAAT SE & Co. KGaA, Einbeck, Germany
[3]Animal Breeding and Genetics Group, Department of Animal Sciences, University of Goettingen, Germany
[4]Heritable Ag, Mountain View, California, USA

Genomic prediction (GP) is a central component of modern plant breeding, enabling the early selection of superior individuals based on genomic data. Classical GP models, such as genomic best linear unbiased prediction (GBLUP), operate within the data modeling culture and typically assume additive genetic effects and linearity, which are limitations that hinder their performance in hybrid breeding where non-additive effects, like dominance and epistasis, play a key role. In contrast, machine learning (ML) models from the algorithmic modeling culture can model complex non-linear interactions but often lack biological grounding and interpretability. To bridge these paradigms, we propose 2NPLGBM, a hybrid framework that integrates quantitative genetics with ML. This method introduces a two-matrix (2NP) genotype representation by concatenating additive and dominance matrix representations, which serves as input to a Light Gradient Boosting Machine (LGBM). We benchmarked 2NPLGBM against GBLUP with additive effects (GBLUP_ADD) and GBLUP with additive and dominance effects (GBLUP_ADDOM) using hybrid maize data across four agronomic traits: grain yield, plant height, days to silking, and days to anthesis. Evaluation under two temporal validation schemes, Leave-One-Year-Out (LOYO) and Rolling Window (RW), revealed that 2NPLGBM consistently outperformed baseline models in both predictive accuracy and selection efficiency under LOYO, with accuracy improvements of 2-14% and selection efficiency gains of 11-17% across all traits. Under the RW scheme, GBLUP_ADDOM outperformed 2NPLGBM in predictive accuracy; however, 2NPLGBM maintained higher selection efficiency, with gains ranging from 2% to 13%. These results demonstrate the value of combining biologically informed genotype encodings with machine learning to enhance predictive performance and selection outcomes in hybrid breeding.

14. Sparse testcrossing for early-stage genomic prediction of general combining ability to increase genetic gain in maize hybrid breeding programs

David O. González-Diéguez[1,2], *Gary N. Atlin*[5], *Yoseph Beyene*[3], *Dagne Wegary*[4], *Dorcus C. Gemenet*[1,3], *Christian R. Werner*[1,2]

[1]Breeding Innovation and Modernization, Consultative Group of International Agricultural Research (CGIAR), Texcoco, Mexico
[2]International Maize and Wheat Improvement Center (CIMMYT), Texcoco, Mexico
[3]International Maize and Wheat Improvement Center (CIMMYT), Nairobi, Kenya
[4]International Maize and Wheat Improvement Center (CIMMYT), Harare, Zimbabwe
[5]Bill & Melinda Gates Foundation, Seattle, WA 98109, USA

Maize hybrid breeding programs aim to develop new hybrid varieties by crossing genetically distinct parents from different heterotic pools, exploiting heterosis for improved performance. The programs typically consist of two main components: population improvement and product development. The population improvement component aims to enhance the heterotic pools through reciprocal recurrent selection based on general combining ability (GCA). However, especially in the early stages of testing, evaluating large numbers of hybrid combinations to estimate GCA is impractical due to considerable logistical challenges and costs. Therefore, breeders often evaluate the initial population of selection candidates using only a single tester to narrow down the candidate pool before further evaluation. Using a single tester, however, may not adequately represent the heterotic pool, leading to inaccurate GCA estimates and suboptimal selection decisions. To address this, we propose sparse testcrossing for early-stage testing, where subsets of candidate genotypes are testcrossed with different testers, connected through a genomic relationship matrix. We conducted stochastic simulations to compare various sparse testcrossing designs with a conventional testcross strategy using a single tester over 15 cycles of reciprocal recurrent genomic selection. Our results show that using a few testers sparsely distributed among full-sibs, sparse testcrossing offers breeders a practical balance between simple testcross designs, resource efficiency, and increased prediction accuracy for GCA, ultimately leading to increased rates of genetic gain.

15. Harnessing genetic diversity from the wider cultivated gene pool to advance trait analysis and breeding in potato

Karen McLean[1], *Mads Sønderkær*[2], *Glenn Bryan*[1], *Sanjeev Kumar Sharma*[1]

[1]James Hutton Institute, Invergowrie, Dundee, DD2 5DA, Scotland, United Kingdom
[2]Aalborg University, Fredrik Bajers Vej 7H, 9220 Aalborg, Denmark

Despite its importance as a global food security crop, potato (*Solanum tuberosum* L.) has seen only modest genetic gains over the past century, and only a limited gene pool has been exploited for breeding. In this study, we harnessed a broader spectrum of cultivated potato diversity, incorporating genotypes from historical and modern cultivars, alongside accessions from the wider cultivated gene pool, including *S. tuberosum* subsp. *tuberosum*, subsp. *andigena*, subsp. *phureja*, and the wild progenitor *S. brevicaule*.

Using our in-house, state-of-the-art whole-exome sequencing platform - targeting all known potato genes, including resistance and microRNA loci - we performed genome-wide scans that identified 10M variants. From these, a set of 90k high-impact SNPs - prioritised through rigorous filtering and spanning the entire gene space - was selected to design a novel gene-based Single-Primer Enrichment Technology (SPET) for potato.

We validated the SPET platform through genetic studies focused on key agronomic, processing, and disease resistance traits. Our findings highlight the power of this approach in tapping unexplored genetic diversity, capturing the majority of the trait genetic variance and maximising novel QTL discovery with higher precision and resolution. The genic origin of SPET markers further enhances their potential to convert QTL-SNPs into effective, trait-linked markers for breeding.

Insights from this research - including temporal trends in cultivated potato diversity, selection footprints, and the genetic architecture underlying key traits - will be presented, highlighting the utility of comprehensive genomic tools to unlock underutilized genetic diversity and to enable more precise, data-informed breeding strategies in potato.

16. Study of the impact of genome editing in a perennial species breeding program through simulations

Xabi Cazenave[1,2], *Jérôme Bartholomé*[3,4,5], *Mathieu Tiret*[6], *Alain Charcosset*[2], *Laurence Moreau*[2], *Leopoldo Sanchez*[1]

[1] INRAE, ONF, BioForA, UMR 0588, 45075 Orléans, France
[2] INRAE, CNRS, AgroParisTech, GQE–Le Moulon, Université Paris-Saclay, Gif-sur-Yvette, France
[3] CIRAD, UMR AGAP, Cali, Colombia
[4] AGAP, Univ Montpellier, CIRAD, INRA, Montpellier SupAgro, Montpellier, France
[5] Crops for Nutrition and Health, International Center for Tropical Agriculture (CIAT), Cali, Colombia
[6] IGEPP, INRAE, Institut Agro, Univ Rennes, Le Rheu, France

Genome editing has been shown to have the potential to improve quantitative traits. This approach could especially be beneficial for perennial tree species which generally have a long breeding cycle. To assess the potential of genome editing for such species, we simulated a poplar breeding program over 10 generations based on recurrent genomic selection to investigate the impact of (i) the number of edited parents (from one to every parent) and (ii) the number of edited QTLs (from one to twenty). Genetic gain was higher by up to 38% than a program with no edition when at least 5 QTLs were edited, and the more parents edited the higher the obtained genetic gain. When most or all the parents were edited, genome editing had no impact on genetic diversity but editing only a few parents led to higher levels of inbreeding and a higher loss of genetic diversity due to the greater contribution of these parents at each generation. When further optimizing the mating plan by using Optimal Cross Selection (OCS), simulations showed that the same levels of genetic gain as a program with no edition could be achieved while maintaining approximately 25% more genetic variance and reducing inbreeding, especially when only a few parents were edited.

17. Robustness evaluation of machine learning models in genomic prediction

Vanda M. Lourenço[1], *Hans-Peter Piepho*[2], *Joseph O. Ogutu*[2]

[1]Department of Mathematics and NOVA Math, NOVA FCT, NOVA University of Lisbon, Portugal
[2]Biostatistics Unit, University of Hohenheim, Germany

Accurate genomic prediction (GP) of breeding values is essential in modern plant and animal breeding programs. GP relies on thousands of molecular markers distributed across the genome, requiring computational methods capable of handling high-dimensional data. In this context, machine learning (ML) has emerged as a powerful framework due to its flexibility and ability to model complex genetic architectures.

While many studies have compared the predictive performance of individual ML algorithms, few have offered broader evaluations across diverse methodological approaches—particularly with respect to robustness under data contamination. Yet, in practical breeding applications, data quality is often imperfect, and accuracy depends not only on model fit but also on a method's ability to perform reliably in the presence of noise.

This study addresses these gaps by evaluating the predictive accuracy and robustness of a range of supervised ML methods. Using simulated data from an animal breeding population, we assess performance across varying levels of data contamination, focusing on both prediction accuracy and error metrics.

The results offer new insights into the relative strengths and limitations of different ML approaches under realistic conditions. These findings provide practical guidance for selecting robust and effective methods for genomic prediction in breeding applications.

18. Types of plant breeding paper

Rex Bernardo

University of Minnesota, Minnesota, USA

Several years ago, a webcomic entitled “Types of Scientific Paper” was posted in xkcd.com. In this webcomic, author Randall Munroe “satirizes scientific papers by claiming they all (or at least the large majority) fall into only a small number of categories, which he describes with somewhat humorous generalized titles” such as “My colleague is wrong and I can finally prove it.” This original comic has spawned specialized versions in different fields including epidemiology, history, neuroscience, plant science, and psychiatry. This poster presentation unveils a plant breeding version (e.g., “We characterized genetic variation in bad germplasm”) that allows us to ponder the current foci of plant breeding papers and reflect on areas of investigation that need more attention.

19. Co-evolutionary analysis for mining functional genes from plant genomes

Shang Gao, *Huihui Li*

Co-evolutionary analysis through phylogenetic profiling (PP) offers a powerful approach to decipher plant gene functions by detecting conserved presence/absence patterns across species. This method is particularly valuable for non-model plants, where traditional genetic tools are limited, and can bridge evolutionary adaptation with modern breeding by identifying stress-resilience genes from wild relatives. By analyzing evolutionary patterns across 365 plant species, we developed a framework enabled accurate identification of salt stress-related genes in the halophyte *Spartina alterniflora*, achieving over 80% prediction accuracy in linking co-evolution patterns to functional conservation. Our approach successfully identified five key ion transporters (including Na^+/H^+ exchangers and K^+ transporters) that were experimentally validated to mediate Na^+ uptake under salt stress conditions, as well as unique R3-MYB orthologs (SaCPC1/2) potentially involved in salt gland development through their co-evolution with circadian rhythm genes. These findings highlight how co-evolutionary analysis can effectively prioritize targeted genes by revealing conserved stress-adaptive modules.

20. Identifying stable and high-yielding oil palm genotypes in a long-term single-site breeding trial

Mohd Ibnur Syawal Zakaria, *Syafeqa Abdul Hamid*, *Sheh May Tam*

SD Guthrie Research Sdn Bhd

Better understanding of stability in oil palm yield performance is crucial for sustainable production, particularly under environmental fluctuations in face of climate change. This study evaluates genotype performance using 15 years of single-location yield recording data via stability analysis in the METAN R package. Traits including Fresh Fruit Bunch (FFB), Bunch Number (BNO), and Bunch Weight (BWT) were assessed using AMMI, GGE, WAAS/WAASB, and multi-trait selection indices (FAI-BLUP, MTSI) with environment defined as year. Significant genotype-by-year interactions were observed, justifying stability modelling. Results showed that G11 excelled during early years (YR1–YR8), while G10, G12, and G13 performed best in mid-to-late years (YR9–YR15). BNO was more environmentally sensitive than BWT, suggesting different breeding priorities per trait. Top performing genotypes were consistent across models, with G10, G12, and G13 showing both yield potential and trait stability. These findings support the use of long-term, single-site trials for in depth knowledge gain and practical breeding decisions, especially where multi-location resources are limited. METAN offers oil palm breeders an accessible, powerful tool to identify stable (climate-resilient), high-performing planting materials.

21. Barley breeding under long days: Reducing speed breeding energy use and unlocking flowering plasticity for climate resilience

Nicola Rossi[1,2], *Rajiv Sharma*[2], *Wayne Powell*[2]

[1]UoE
[2]SRUC

Flowering time is a critical determinant of yield in barley (*Hordeum vulgare*), a long-day species that flowers once photoperiod requirements are met. Speed breeding (SB) protocols typically use a 22-hour photoperiod to accelerate development, but the minimum effective day length remains undefined. Reducing photoperiod could lower energy costs while maintaining flowering acceleration. At the same time, climate change is shortening growing seasons, increasing the need for allelic combinations that promote earlier flowering without yield penalties.

To address these challenges, we used Bayesian hierarchical modelling to study the flowering time response to increasing day length in relation to allelic variation at PPD-H1, ELF3, and PHYC. Flowering was recorded under five photoperiods (16h–24h) in Near Isogenic Lines (NILs) carrying elf3 (eam8.w) or PhyC-e alleles, and in Recombinant Inbred Lines (RILs) from the HEB-25 population with factorial combinations of wild (*Hordeum spontaneum*, Hsp) and domesticated (*Hordeum vulgare*, Hv) ELF3 and PPD-H1 alleles. Gene expression analysis supported the phenotypic trends. We found that: (i) lines that harbour PPD-H1Hv responded to a 20h threshold, suggesting energy savings over standard SB protocols, while PPD-H1Hsp lines did not respond in this range, consistent with a known 16h threshold. (ii) ELF3Hsp in a PPD-H1Hv background reduced intrinsic earliness without affecting plasticity. (iii) PhyC-e reduced plasticity but not earliness.

Our findings suggest that a 16–20-hour SB protocol can be tailored to genotype, significantly reducing light use. Moreover, ELF3Hsp and PhyC-e alleles in a ppd-H1 background offer valuable tools for breeding barley lines adapted to future climatic conditions.

22. An interpretable machine learning-based alternative to genome-wide association studies (GWAS) and its application in a wild population

Gard W. Gravdal[1,3], *Henrik Jensen*[2,3], *Hamish A. Burnett*[2,3], *Stefanie Muff*[1,3]

[1]Department of Mathematical Sciences, NTNU, Trondheim, Norway
[2]Department of Biology, NTNU, Trondheim, Norway
[3]Gjærevoll centre, NTNU, Trondheim, Norway

The identification of genetic markers associated with phenotypic traits is important in evolutionary biology, medicine, and breeding, but studies in wild populations remain rare. Traditional genome-wide association studies (GWAS) have limitations for complex traits and rely on p-values, which combine effect size and uncertainty, making them unreliable measures of variable importance.

We investigated alternative measures of variable importance for identifying associations between Single Nucleotide Polymorphisms (SNPs) and phenotypic traits. "Shapley values" were originally constructed for regression models to capture each variable's contribution to the variance explained. We employed machine learning models including boosted regression trees (XGBoost), and estimated variable importance using Shapley Additive Explanation (SHAP) values. The approach was tested on genomic data with $\sim$70,000 SNP markers and phenotypic measurements of $\sim$4000 wild house sparrows (*Passer domesticus*) from a metapopulation off the Helgeland coast in northern Norway, analysing body mass, tarsus length and wing length. For comparison, we performed a classical GWAS using linear mixed models. We also assessed the methods on simulated phenotypes based on additive genetic models for a range of genetic architectures.

SHAP results give complementary information to GWAS, with little overlap in SNPs identified. GWAS identified no SNPs passing the Bonferroni threshold for the three highly polygenic traits, illustrating its limited power under realistic architectures and sample sizes. This is consistent with our simulations, where both methods failed for highly polygenic traits but detected SNPs under more oligogenic scenarios. In addition, SHAP values estimate phenotypic variance explained per SNP, offering insights into genetic architecture.

23. The impact of increased recombination on breeding programs: Insights from simulations

Boyny Zsa Zsa, Lester Nicholas, Massel Karen, Powell Owen, Snowdon Rod, Weber Sven

Recombination shuffles alleles during meiosis, driving genetic diversity and shaping the outcomes of breeding programs. By breaking the physical links between loci, recombination facilitates the creation of new allelic combinations that are easier to select for and increase genetic gain. Increasing recombination rate by methods such as genome editing has become a goal for accelerating breeding. However, the effect of increased recombination rate on a population scale on breeding programs is not fully understood. We therefore carried out simulations to determine the effect on a breeding program using both phenotypic and genomic selection. We focused on how heritability, number of quantitative trait loci, recombination rate increase factor, marker density and training frequency affect breeding success. We also tested whether it is possible to use historic training sets without recombination rate change and merge the pre- and post-recombination populations to improve prediction accuracy and genetic gain in genomic selection. We found that increasing recombination is particularly beneficial for highly quantitative traits with low heritability. However, with genomic selection, increasing recombination requires a higher training frequency as well as an increased marker density to accelerate superiority over phenotypic selection in terms of genetic gain. Furthermore, our simulations show that maintenance of old training sets and merging of training sets with different recombination rate is possible, but a decrease in prediction accuracy is expected, favouring frequent training and high marker density under increased recombination rates.

24. Permutation-based GWAS in raspberry

Philip Greenspoon[1], *Julie Graham*[2], *Brezo Mateos*[2], *Susan McCallum*[2], *Valeria Montano*[1], *Manon Verdier*[1],

[1] Biomathematics and Statistics Scotland (BioSS)
[2] James Hutton Institute

Identifying the genetic basis of relevant traits through genome-wide association mapping (GWAS) can improve the efficiency of a phenotypic-based breeding program, as discovered genetic markers can inform marker assisted selection (MAS). Typically, GWAS is performed by repeated statistical tests across genetic markers, which assumes independence among genetic markers and that the model residuals follow a normal distribution. Either or both assumptions are often unmet. Independence among markers is usually invalid because genetic markers exhibit linkage disequilibrium, and wrongly assuming it can lead to excessive false negatives. Normality of residuals is also likely to be violated, particularly for traits that are not continuous, and incorrectly assuming it can lead to excessive false positives. In this work, we perform GWAS on raspberries, using a recently released permutation-based tool, permGWAS2, which does not depend on these assumptions and is computationally efficient. Raspberries are a high value horticultural crop with about 16 thousand tonnes produced in the UK in 2023. Confronting climate change and market competition there is a need to continuously produce improved raspberry cultivars. We found new markers for plant development timing traits and fruit quality traits. Identifying genetic markers for developmental timing traits is valuable in the face of climate change, because they can be used in MAS to help develop cultivars synchronised with shifting timing of seasonal changes, while markers for fruit traits are helpful for maintaining market value.

25. Building foundations for ‘Ensembl plant populations’

Bedford, J.A.[1], *Love, B.*[1], *Wright, T.I.C.*[1], *Falola, O.*[2], *Connell, J.*[1], *Flint, B.*[2], *Harrison, P.*[2], *Saraf, S.*[2], *Alvarez-Jarreta, J.*[2], *Giorgetti, S.*[2], *Lodha, D.*[2], *Percival-Alwyn, L.*[1], *Naamati, G.*[2], *Cockram, J.*[1], *Dyer, S.*[2]

[1] Niab, Park Farm, Villa Road, Histon, Cambridgeshire, CB24 9NZ, UK
[2] EMBL-EBI, Wellcome Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK

The United Kingdom has strong capabilities in plant research, with extensive expertise in germplasm resource development, trait discovery platforms, and genomics. However, integrating these areas has historically been challenging, requiring significant technical expertise. The ‘Ensembl Plant Populations’ initiative is a collaborative effort between Ensembl and Niab, that has been funded by the BBSRC to address these challenges head-on. This open-source online platform will provide researchers with the necessary tools for robust genome-wide association analyses and candidate gene assessment for 12 already genotypically defined and publicly available plant populations.

Focusing on seven key plant species - Arabidopsis, barley, oat, oilseed rape, rice, tomato, and wheat - the project combines carefully curated genotype data with genetic analysis pipelines to ensure statistical integrity. Researchers will be able to access seed from the relevant Germplasm Resource Unit, make their measurements and observations, before uploading their phenotypic data to ‘Ensembl Plant Populations’. This platform will then perform association studies and candidate gene assessment and the user will receive expert guidance through an interactive visualisation tool to identify key genetic components, and directly link to the existing Ensembl Plants infrastructure, enabling further bioinformatic exploration of resulting genetic regions of interest. This project will enable a significant step forward in the community’s ability to explore the genetic basis of important plant traits, in a number of key species, streamlining the research process and fostering collaborative discovery.

26. Predicting novel genotypes in untested environments using large multi-environment datasets across species

Vincent Garin, *Simon Rio*, *Julien Frouin*, *Alice Boizet*, *Marion Buffard*, *Stéphanie Sidibe-Bocs*, *Carlos Viquez-Zamora*

CIRAD

We evaluated statistical models and sampling strategies to predict the genetic values of new (untested) genotypes in new (unobserved) environments using three medium- to large-scale multi-environment trial networks. We used the maize G2F dataset, which includes 5,000 genotypes and 272 sites (mostly in the US); the sorghum WCA-BCNAM dataset, composed of 3,905 genotypes evaluated in 11 environments (Mali); and the rice PRAY association panel, composed of 281 genotypes evaluated in 10 environments across the globe (Colombia, Senegal, India, China, and the Philippines).

To predict genotypic values, we compared linear mixed models (unstructured model, reaction norm model, and factorial regression models) with standard machine learning algorithms such as light gradient boosting machine, neural networks, random forests, and support vector regression. The linear mixed models generally achieved higher prediction accuracy than the machine learning approaches. In some cases, ensemble approaches combining methods from both classes showed improved predictive performance.

We constructed training sets with the aim of maximizing either the coverage of the environment set or the genotype set. In datasets where the genotype distribution is balanced across environments, sampling strategies that maximize genotype coverage tend to yield better results. In datasets where genotype distribution is more unbalanced across environments, it is preferable to include a larger number of environments in the training set.

27. Strategic crossing to improve genetic potential in soybean breeding

Kengo Sakurai[1], *Yusuke Toda*[2], *Minoru Inamori*[1], *Kosuke Hamazaki*[3], *Hisashi Tsujimoto*[4], *Akito Kaga*[5], *Hiroyoshi Iwata*[1]

[1]Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo 113-8657, Japan
[2]Institute for Agro-Environmental Sciences, National Agriculture and Food Research Organization, Tsukuba 305-8604, Japan
[3]Molecular Informatics Team, RIKEN Center for Advanced Intelligence Project (AIP), RIKEN, Chiba 277-0871, Japan
[4]Arid Land Research Center, Tottori University, Tottori 680-0001, Japan
[5]Institute of Crop Science, National Agriculture and Food Research Organization, Tsukuba 305-8518, Japan

In breeding programs, balancing genetic gain with the maintenance of genetic potential is essential. While selecting superior genotypes improves genetic gain, it often reduces the overall genetic potential of the population. The Optimal Population Value (OPV) is a criterion used to evaluate genetic potential. Despite its importance, few studies have aimed to improve OPV. In this study, we aimed to enhance the OPV of a soybean breeding population through targeted crossing.

We used 186 recombinant inbred lines (RILs) derived from a cross between ‘Enrei’ and ‘PK 73-54’. From these, eight lines were selected to produce 14 F_2 lines, each with approximately 64 progeny, totaling 897 individuals (C1F2). These F_2 lines produced F_4 lines (C1F4), comprising 820 genotypes. Genotypic data (524 SNP markers) were collected for the RILs, C1F2, C1F4 populations. Additionally, biomass data were collected for 186 genotypes in the RILs, 235 in C1F2, 208 in C1F4. A genomic prediction model was built and estimated marker effects using biomass and genotypic data. OPV was calculated for all pairwise combinations among the 14 F_2 lines in the C1F2 population. Based on these values, 28 crosses were selected using each F_2 line twice to improve OPV, yielding 382 progeny (C2F2), with approximately 16 individuals per cross. We also collected genotypic and biomass data of C2F2 population.

The C2F2 population exhibited the highest OPV among the four populations, exceeding C1F2 and C1F4 by 3% and 5%, respectively. These findings demonstrate that OPV can be improved through strategic crossing.

28. Genome-wide association study of agronomical and nutritional traits in oat using a recurrent selection population with *Avena sterilis* introgressions

Kai Ilves[1], *Min Lin*[1], *Espen Sørensen*[2], *Therese Birkeland Fossøy*[3], *Hilde Halland*[4], *Sigridur Dalmannsdottir*[4], *Hrannar Smari Hilmarsson*[5], *Lidija Bitz*[6], *Oliver Bitz*[6], *Hanna Haikka*[7], *Pernilla Vallenback*[8], *Fiona Doohan*[9], *Cathal McCabe*[9], *Atikur Rahman*[1]0, *Lucia Gutierrez*[1]1, *Juho Hautsalo*[6], *Morten Lillemo*[1]

[1]Norwegian University of Life Sciences, Department of Plant Sciences, Ås, Norway
[2]Graminor, Ridabu, Norway
[3]Norwegian Institute for Bioeconomy Research, Steinkjer, Norway
[4]Norwegian Institute for Bioeconomy Research, Tromsø, Norway
[5]Agricultural University of Iceland, Hvanneyri, Iceland
[6]Natural Resources Institute Finland, Rovaniemi, Finland
[7]Boreal Plant Breeding, Jokioinen, Finland
[8]Lantmännen Lantbruk, Svalöv, Sweden [9]University College Dublin, School of Biology and Environmental Science, Dublin, Ireland
[1]0Teagasc, Oak Park, Carlow, Ireland
[1]1University of Wisconsin – Madison, Department of Agronomy, Madison, Wisconsin, USA

Oat cultivation in Nordic countries has expanded due to novel food uses, its resilience in unfavorable climate, and good fit for organic farming. The homogeneous Nordic oats withstand the harsh climate of the northern periphery, but their nutritional composition is less optimal for producing new oat-based foods from local cultivars. To solve this, we characterize an oat recurrent selection population consisting of 20 parents of Nordic and North American origin, including introgressions from *Avena sterilis*, alongside their progenies from 6 recurrent selection cycles, for their potential as diversity increasing breeding material. This population has been genotyped by a 7k SNP chip and genotyping-by-sequencing, and as a part of the OatFrontiers (2023-2026) project, grown in 11 locations across 6 countries for phenotyping different agronomic and harvest traits.

Days to heading, maturity, and plant height manifested relatively stable genotype effects across environments, whereas the yield, fat, and protein content showed significant GxE effects. GWAS with BLINK model revealed significant markers for earliness on chromosome 7D, yield on 1D, 2D and 5D, height on 5C, 7A and 7D, lodging on 1D and 6A, fat on 4C and 6A, and protein content on 2C, 2D,3A and 7D. Chromosome 7D contained highly significant markers for earliness in the 469 Mbp region (Asativa_sang.v1.1) in almost all locations. Regardless, the earliness alleles of these markers correlated negatively with yield. Further analysis is needed for confirming the significant earliness loci without harvest trait penalties to select the best crossing parents for further breeding.

29. Genotype-by-environment interactions in Norwegian barley: insights from a decade of multi-location trials.

Min Lin[1], *Shirin Mohammadi*[1], *Nora Røhnebæk Aasen*[2], *Silius Mortensønn Vandeskog*[2], *Alex Lenkoski*[2], *Maria Thorkildsen*[3], *Morten Lillemo*[1]

[1]Department of Plant Sciences, Norwegian University of Life Sciences, Post Box 5003, NO-1432 ÅS, Norway
[2]Statistical Analysis and Machine Learning, Norwegian Computing Center, Oslo, Norway
[3]Division of Food Production and Society Grain and Forage Seed Agronomy, Norwegian institute of bioeconomy research

Barley (*Hordeum vulgare* L.) is the most widely cultivated cereal crop in Norway and plays a pivotal role in the country’s agricultural sector. Due to geographical constraints, barley production is primarily concentrated in three major regions: Mid Norway, Northeast Norway, and Southeast Norway. Each year, official variety trials are conducted in these regions to evaluate the performance of candidate genotypes prior to their release and registration as commercial cultivars. Understanding genotype-by-environment (GxE) interactions, identifying optimal testing locations, and best performing cultivars for each region are therefore essential for breeding programs, seed companies, extension service and farmers. In this study, we analyzed ten years of data (2014-2024) from official yield trials conducted across different locations in Norway, involving approximately 80 barley genotypes in total, with an average of 26 genotypes tested annually across 13 different environments. To assess GxE interactions, we employed both the AMMI (Additive Main Effects and Multiplicative Interaction) and GGE (Genotype and Genotype-by-Environment interaction) models. Additionally, we incorporated climate data from the ERA5 reanalysis dataset for all trial locations over the same period. One of the main objectives of this study is to elucidate how climatic variability influences GxE interactions of barley yield, and to identify both the most regionally adapted cultivars and the most stable performers across diverse environments in Norway. Furthermore, the integration of climate data enables us to predict genotype performance under specific weather scenarios, thereby enhancing preparedness for future climate change

30. Improving genomic prediction in wheat with random regression models for environmental covariates.

Rishap Dhakal[1], *Guillermo Sniadower*[2], *Paula Silva*[3], *Betina Lado*[2], *Pablo Sandro*[1], *Inés Rebollo*[4], *Martin Quincke*[1], *Pablo González Barrios*[2], *Lucia Gutiérrez*[1]

[1]Department of Plant and Agroecosystem Sciences, University of Wisconsin-Madison, 1575 Linden Drive, Madison, WI 53706, USA
[2]Facultad de Agronomía, Universidad de la República, Av. Garzón, 780, Montevideo 12900, Uruguay
[3]Instituto Nacional de Investigación Agropecuaria (INIA), 70000 Colonia, Uruguay
[4]Department of Agronomy and Plant Genetics, University of Minnesota, Saint Paul, USA

Genotype-by-environment interaction (GEI) creates a challenge for selecting top-performing genotypes. Mixed models for genomic prediction that model the variance-covariance structures based on crop performance, such as factor analytic (FA) model, generally have high predictive ability (PA) but could only predict the performance in tested environments. Random regression models (RRM) that include environmental covariates (EC) could be a better option, as they can predict the performance of genotypes in untested and also tested environments when crop performance data is not available to fit models like FA. Our main goal was to evaluate genomic prediction models that integrate GEI and EC in a large dataset. We used a yield dataset consisting of 1683 genotypes across 79 trials with 71 environments in 11 years, from six locations in the National Wheat Breeding Program (WBP) of Uruguay. We hypothesize that RRM with EC improves predictive ability for both tested and untested environments. We fitted different genomic prediction models, including GBLUP, GBLUP G×E FA2, and RRM models, and tested their PA using different schemes: CV1 (untested genotype), CV2 (tested genotypes) in tested environment, and CV0 (genotype in untested environments). RRM models outperformed GBLUP by 124% in CV1 and 70% in CV2. For CV1, PA was very close to that of FA2. For CV0, the RRM model had higher predictive ability than GBLUP in more than 44 out of 71 environments. Thus, RRM models could be effectively utilized in breeding programs for predicting genotype performance considering GEI and can result in higher response to selection.

31. QTL mapping and candidate genes associated with common bean resistance to root-knot nematode (*Meloidogyne incognita*)

Bruna Marques Moreno[1], *Lílian Beatriz Januario Bibiano*[1], *Talissa Oliveira Floriani Zimermman Souza*[2], *Antonio Augusto Franco Garcia(*[1], *Guilherme da Silva Pereira*[3], *Maria Lucia Carneiro Vieira*[1]

[1]Luiz de Queiroz College of Agriculture
[2]Illinois State University
[3]Federal University of Viçosa

The common bean (*Phaseolus vulgaris*, 2n = 22) is a staple crop worldwide, yet its productivity is severely affected by root-knot nematodes (*Meloidogyne incognita*), for which no complete resistance has been identified. This study aimed to uncover genomic regions associated with nematode resistance using an F2:3 biparental population derived from a cross between two Mesoamerican black bean cultivars: Puebla-152 (moderately resistant) and Jamapa-CNF-1671 (susceptible). Plants were phenotyped under controlled conditions in a growth chamber following inoculation with 1,500 second-stage juveniles (J2), with galling index (GI) and number of egg masses (NEM) assessed after 30 days. Genotyping was performed using SNP markers from Genotyping-by-Sequencing (GBS), and a linkage map was constructed using the OneMap package. QTL mapping was carried out using Composite Interval Mapping in R (r/qtl), and broad-sense heritability was estimated using the Cullis method. A major QTL for NEM was identified on chromosome Pv01, with a peak LOD score of 6.7, additive and dominance effects of -1.1/-6.68 and explaining 11.44% of phenotypic variation. The heritability estimate for NEM was 34.1%. Notably, this QTL region overlaps with loci previously associated with disease resistance and root architecture. Within the QTL interval, several candidate genes were identified, including transcription factors, kinases, and defense-related proteins. Candidates include Phvul.001G038100 (pectin methylesterase inhibitor), Phvul.001G038800 (cytokinin dehydrogenase), and Phvul.001G043800 (F-box protein). These findings contribute to the understanding of resistance mechanisms and propose candidate targets for subsequent validation and potential application in marker-assisted breeding for durable nematode resistance in common bean.

32. Implementation of genomic selection in *Miscanthus sinensis* for ecosystem services

Manuel Derrien, *Séverine Monnot*, *Maryse Hulmel*

Miscanthus is becoming increasingly popular in France, with over 1,000 new hectares being planted each year. Miscanthus is a perennial crop that reproduces clonally and can produce for about twenty years without chemical inputs, fertilizers, irrigation or any other interventions than harvest from the grower. As such, it is considered as a model species for studying complex traits related to ecosystem services, such as nitrogen remobilization and water quality. However, only one interspecific clone of *Miscanthus* x *giganteus*, is cultivated worldwide, which could weaken the sector in the event of severe weather event or the emergence of pathogens. The INRAE BioEcoAgro team aims to expand the varietal range by working on *Miscanthus sinensis*, one of the two parental species. This species, being wild, interspecific, perennial, clonal, and associated with the measurement of complex traits, presents interesting challenges for adapting conventional selection models. In this context, we are developing genomic and phenomic selection models that incorporate the age of the crop using a stagger-start design. The results are promising for traits such as biomass, and genomic selection will help to reduce the selection cycles of Miscanthus from four years to one year.

33. Maximizing G×E value in forestry breeding using a three-stage approach

Mason Chizk, John Moore, Christine Te Riini, Yue Lin, Jude Sise, Bruno Santos

In forestry breeding, genotype-by-environment (G×E) interactions play a major role in determining clonal performance across landscapes. While mixed models such as single-step GBLUP (ssGBLUP) are well established for modeling G×E, breeders also face a practical constraint: how many clones should be deployed in large scale forestry operations to balance the value of environment-specific gains with the complexity of managing diverse portfolios? We present a three-stage strategy to optimize clonal deployment under this trade-off. In Stage I, we used an environmental ssGBLUP model to predict clonal performance across environments. In Stage II, we reduced complexity by isolating top-performing clones and clustering them based on their predicted G×E response profiles using k-means, across a range of k values representing potential portfolio sizes. In Stage III, we performed a full combinatorial search to identify the optimal combination of clones across clusters—assuming each is deployed to its best-performing environment—thus maximizing total predicted forest value. This framework enables breeders to quantify the trade-off between deployment simplicity and overall productivity. We show how gains in forest value plateau as more clones are added. Unlike approaches that rely on predefined ecological zones, our strategy allows the data to define functional deployment groups, revealing novel patterns of clonal adaptation across environments.

34. From fields to fjords: a framework for targeting resilience by simulating longitudinal growth traits in plants and aquaculture

Duncan Henderson, *Smaragda Tsairidou*, *Dominic Waters*, *Daniel Tolhurst*

Genotype-by-environment (GxE) interaction can reduce the efficacy of plant, terrestrial livestock and aquaculture breeding programs, especially when reranking of genotypes occurs between environments. While GxE interaction is well-integrated in plant breeding, its potential in aquaculture remains relatively underutilized. Like plants, aquaculture performance is heavily reliant on the environmental conditions they are grown in, so that differences between broodstock and commercial production settings (e.g., open-sea farms/cages) impede effective selection, as the top-performing families may underperform in commercial farm environments. This parallels challenges seen during field evaluations in plant breeding and the need to consider a breeder's target population of environments.

This work presents a framework for simulating longitudinal growth traits in aquaculture by adapting methods from plant breeding to evaluate the benefits of incorporating environmental resilience into the breeding objective. We simulate Atlantic salmon genotypes under realistic breeding scenarios using AlphaSimR, integrating 20 years of real-world climate data for four key salmon farming locations: Scotland, Norway, Tasmania and Chile. We model commercially important performance traits such as growth, alongside resilience-related traits such as environmental sensitivity and stress recovery over time.

This framework will highlight the importance of incorporating environmental resilience into breeding schemes by leveraging GxE interaction present in aquaculture to buffer against the effects of climate change, enabling the selection for genotypes suited for diverse production environments. We will also show how the framework has applications to plant breeding settings, particularly with the advent of high-throughput and hyper-spectral longitudinal data recorded over the growing season.

35. Genotype and phenotype encryption optimised for federated quantitative genetics

Arun Isaac[1], *Hao Cheng*[2], *Richard Mott*[1]

[1]University College London Genetics Institute
[2]UC Davis Department of Animal Science

We have developed a framework for encrypting genome and phenotype data that is well suited to many quantitative genetics analyses. The key idea is to apply a random orthogonal transformation to the genotype dosage matrix and the matrix of phenotypes and covariates. The usual Gaussian likelihood is invariant under orthogonal transformation, so that all inferences based on the encrypted data are identical to those using the plaintext. This is a form of Homomorphic Encryption, but which differs from that used in other fields in that there is no need to decrypt the data in order to obtain results as they are invariant. The range of analyses that can be accommodated includes (but is not limited to) the standard mixed model using a SNP-based genetic-relationship matrix, and the Bayesian Alphabet models where the likelihood is multiplied by a prior on the SNP effects, and the estimation of heritability (see PMIDs 32327562, 38085098 for details). The framework is ideally suited to federated analyses where a consortium wishes to jointly analyse their data to increase statistical power, but without decryption back to the plaintext data. We have implemented a federated version of the method in a Python package pyhegp (https://github.com/encryption4genetics/pyhegp).

36. Comparing uni-modal vs multi-modal ML modeling of GxE interactions on hybrid maize data

Iñigo Azqueta, *Claudia Serrano Colomé*, *Finn Gaida*

NoMaze

Genomic prediction in plant breeding must balance statistical power with biological realism when modelling genotype-by-environment (G×E) interaction. Linear mixed models for genomic selection traditionally treat the genetic (G), environmental (E) and G×E components as separate fixed or random effects, whereas most modern deep-learning pipelines ingest SNPs and weather jointly and let the network discover G×E interactions implicitly. Here we show that a neural network which preserves the classical separation of training parallel G-, E- and low-rank G×E sub-modules that exchange information only through a final additive layer, achieves the highest predictive accuracy on the public Genomes-to-Fields (G2F) multi-environment maize trials.

The 2025 G2F release comprises more than 180000 plot records from 5 000 hybrids evaluated in 280 unique year-location environments, with dense marker and hourly weather data [1]. In all cases, we evaluate under the same leave-one-year-out schedule that emulates forecasting into unseen environments. Each data modality is handled by its own architecture and hyper-parameter set: an MLP-based encoder with modality-specific latent size and dropout for SNP markers, a fine-tuned attention-based model for weather records, and an autoencoder for the G×E interaction. Preliminary experiments confirm the expected performance ordering linear < uni-modal < multi-modal, while the multi-modal design remains compact because low-rank factorisation limits the G×E parameter count.

Because the model keeps the genetic and environmental streams distinct, its saliency analysis can rank SNP blocks by influence and pinpoint the most critical weather windows aligned with key growth stages. This dual ranking shows which loci matter most and under what stage-specific heat or moisture conditions those loci drive yield.

These findings confirm the theory that keeping the traditional G + E + G×E decomposition inside a deep-learning framework delivers state-of-the-art accuracy and domain-relevant insight without additional data collection or computational cost.

37. For greener whisky: barley's genetic control for agronomic and metabolic adaptation to reduced nitrogen fertilizer inputs

George Terry Epaku[1,2], *Barry Harrison*[3], *Kelly Houston*[1], *Timothy George*[1]

[1]The James Hutton Institute, Invergowrie, Dundee, UK
[2]School of Life Sciences, University of Dundee, Nether gate, Dundee, UK
[3]The Scotch Whisky Research Institute, Edinburgh, UK

We aim at contributing towards developing Nitrogen-Use-Efficient (NUE) malting barley through (i) assessing the genetic component for adaptation to reduced N inputs, (ii) exploring the balance between carbon and N metabolism in barley under such conditions, and (iii) linking these genetic and metabolic profiles. 232 diverse accessions have twice, been treated with reduced and optimum N for malting barley under field conditions. Even with superior agronomic performance (tillering, heights, yield) under optimum N, plants grown under N deprivation bore heavier (thousand grain weight) and plumper (grain width) grains, with associated genetic regions of each identified. Under N deprivation, elite accessions had greater grain yields and total polysaccharide content (TPC) than their non-elite counterparts. However, the non-elite accessions had higher grain protein content than the elite accessions under this treatment. We have embarked on ascertaining the quality of these polysaccharides and protein for malting and distillation basing on amylose, alpha-amylase, and beta-glucan grain contents. We will then select elite and non-elite accessions under N deprivation with (i) high grain yield, (ii) high TPC with larger amylose than amylopectin and beta-glucan proportions, and (iii) acceptable protein content with large proportions of alpha-amylase. These accessions will be studied for barley's metabolic ability to balance carbohydrate and protein constitutions under N deprivation. We intend to link these metabolic adaptations to genetic constitution for information of future NUE barley variety development. Further, these accessions will be assessed for malting quality under N deprivation, after which they will be fermented and distilled for spirit flavor profiling.

38. Rare genetic variation underlies extremely low trigonelline levels in maize grain

Michael A. Gore[1], *Marjorie Hanneman*[1], *Xiaowei Li*[1], *Ryokei Tanaka*[1], *Laura E. Tibbs-Cortes*[2], *Nolan Bornowski*[3], *Joshua C. Wood*[4], *Di Wu*[1], *Maria Magallanes-Lundback*[5], *John P. Hamilton*[4,6], *Nicholas T. Deason*[5], *Gregory R. Schoenbaum*[2], *Jianming Yu*[2], *C. Robin Buell*[4,6,7], *Dean Dellapenna*[5]

[1]Plant Breeding and Genetics Section, School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA
[2]Department of Agronomy, Iowa State University, Ames, IA 50011, USA
[3]Department of Plant Biology, Michigan State University, East Lansing, MI 48824, USA
[4]Center for Applied Genetic Technologies, University of Georgia, Athens, GA 30602, USA
[5]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824, USA
[6]Department of Crop and Soil Sciences, University of Georgia, Athens, GA 30602, USA [7]Institute of Plant Breeding, Genetics and Genomics, University of Georgia, Athens, GA 30602, USA

Vitamin B3 (niacin: nicotinamide and nicotinic acid) is an essential nutrient largely unavailable in maize-based diets, primarily due to the proportion of trigonelline, a methylated plant-active derivative of nicotinic acid. Trigonelline constitutes a major portion of niacin-related compounds in maize grain but does not release bioavailable nicotinic acid during human digestion. Identifying the genes involved in niacin accumulation and its conversion to trigonelline could improve the nutritional value of maize for human consumption. To explore this, we analyzed niacin and trigonelline levels in grain from $\sim$1,700 maize inbred lines and identified three lines with near-zero trigonelline and approximately 50% higher nicotinic acid levels. One of these extreme lines was selected and crossed with a maize inbred line exhibiting average levels of both traits to construct a biparental population of $\sim$250 F2 individuals for genetic mapping of these grain metabolites. This analysis identified a major locus that explained a large portion of the variation in trigonelline concentration. Among the genes in this interval, a potential causal gene—a methyltransferase—contained a $\sim$5-kilobase transposon inserted in its first exon in the parental inbred line with near-zero trigonelline, which may affect its function. A genetic marker for the presence or absence of the transposon insertion perfectly cosegregated with the trigonelline phenotype in the F2 population. Ongoing work aims to confirm the role of the methyltransferase and further uncover the genetic regulation of niacin and trigonelline accumulation, with the aim of guiding future maize breeding efforts to reduce vitamin B3 deficiencies globally.

39. Advancing genomic selection in grapevine: development and analysis of a training population

Valeria De Rosa[1*], *Elisa De Luca*[2], *Luigi Falginella*[2], *Francesco Flagiello*[1], *Laura Zanin*[1], *Rachele Falchi*[1], *Alessandro Vannozzi*[3], *Emanuele De Paoli*[1]

[1]Department of Agricultural, Food, Environmental and Animal Sciences (DI4A), University of Udine, Via delle Scienze 206, 33034 Udine, Italy
[2]VCR Research Center, Vivai Cooperativi Rauscedo, Via Ruggero Forti 4, 33095 San Giorgio della Richinvelda, Italy
[3]Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), University of Padova, Agripolis, Viale dell'Università 16, Legnaro, Italy

Grapevine cultivation faces increasing challenges due to global climate change, including increasingly frequent heatwaves and drought, and accelerated phenological development. This shift critically impacts the entire growing cycle, from budbreak to harvest, leading to greater susceptibility to late frosts in spring and changes in grape quality and composition. Genomic selection (GS) has emerged as a promising tool to address these challenges, enabling the prediction of phenotypic traits using whole-genome information instead of known markers like in traditional marker-assisted selection (MAS). GS is particularly valid for the improvement of complex quantitative traits controlled by multiple, often elusive, genetic loci. However, the effectiveness of GS depends on the development of a training population that reflects the genetic and phenotypic diversity relevant to breeding objectives.

To this end, a half-diallel crossing scheme was performed in spring 2023 using seven selected parental genotypes. This effort resulted in approximately 2100 seedlings from 21 biparental cross combinations of 100 individuals each. Genotyping was performed using double-digest restriction site–associated DNA sequencing (ddRAD-seq), which generated data from over 40k genomic loci and revealed an average of 90k polymorphic sites. In association to genotypic data, hyperspectral imaging with infrared sensors was performed in spring 2025 aiming to detect hidden phenotypic differences in the early stages of plant development that could be exploited for phenomic selection.

The establishment of this new grapevine training population, named OBIETTIVO2100, represents an essential step toward the development of predictive models for the varietal renewal demanded by a changing climate.

40. Uncovering genomic regions controlling root quality traits in cassava implementing various GWAS models

Diana Carolina Solarte Certuche[1], *Gabriel Mamedio de Freitas*[1], *Jean-Luc Jannink*[2], *Eder Jorge de Oliveira*[3], *Antonio Augusto Franco Garcia*[1]

[1]São Paulo University, Genetics Department, Piracicaba, São Paulo, Brazil.
[2]USDA-ARS, Robert W. Holley Center, Ithaca, NY, United States of America.
[3]Embrapa Mandioca e Fruticulture, EMBRAPA, Genetic Resources and Cultivar Development Center, Cruz das almas, Bahia, Brazil

Biofortified varieties of cassava are particularly important for addressing provitamin A deficiency, which can lead to blindness in children and pregnant women in developing countries. Additionally, dry matter is a key component of cassava yield and has various industrial applications. This study aims to help identify candidate genes associated with dry matter and carotenoid traits within Brazilian cassava germplasm. For the genome-wide association mapping, we utilized two datasets: 25,923 SNPs from DarT technology and 27,045 SNPs from the GBS technology genomic dataset. The analysis was conducted using a population of 3,043 individuals from the EMBRAPA-Brazil germplasm bank, which included 10 years of historical phenotypic data. For the GWAS analysis, we employed two models, MLMM and MLM, allowing for effective control of population structure and relatedness among individuals.

Our findings identified five significant genetic variants associated with carotenoid production located on chromosomes 2, 4, and 10. They are linked to genomic regions that code for proteins involved in the lipid metabolic process, activation of protein kinase, hydrolase activity, and carbohydrate metabolism. For dry matter content, two SNPs are located on chromosomes 1 and 10. These variants are found in genes that regulate the expression of proteins involved in cell wall modification, gibberellin regulation, and carbohydrate metabolism; these biochemical processes are essential for carotenoids and dry matter content. Overall, this complex of seven key genes and their associated genetic variants provides valuable insights into the genetic architecture of these traits and a strong basis for genomic selection models and marker-assisted selection for cassava breeding programs.

41. Genetic analysis of a short-term selection experiment with *Saccharina latissima*: lessons learned

Signe Bråtelund, Gunnar Klemetsdal, Tom Ruttink, Jørgen Ødegård, Franz Goecke, Åshild Ergon

The European kelp industry is rapidly expanding, yet the development of high-yielding cultivated material remains limited. We have conducted a short-term selection experiment in *Saccharina latissima* aimed at improving blade length. Wild individuals were initially crossed in a mixed hybridization to generate an F1 population, from which the longest individuals were again crossed to produce F2 progeny.

Parentage assignment revealed unexpected reproductive modes, including selfing and apomixis, resulting in some highly inbred individuals. We also observed substantial variation in reproductive success among parents, indicating strong unintended selection. Using genotyping-by-sequencing (GBS) data, we constructed a genomic relationship matrix (GRM) and applied genomic best linear unbiased prediction (GBLUP) to estimate trait heritabilities and assess inbreeding depression. Highly inbred individuals exhibited significant inbreeding depression for all the studied traits, namely blade length, blade width, stipe length and individual fresh weight. The analyses confirmed that blade length, blade width, and stipe length were significantly heritable. However, despite phenotypic selection for increased blade length, no genetic gain was achieved for this trait.

The insights gained from this selection experiment, including heritability estimates, genetic correlations between traits and inbreeding effects, are being integrated into simulation models to guide the design of future breeding strategies for *S. latissima*. In future experiments guided by these simulations, we aim to determine optimal gametophyte crossing schemes to obtain high-yielding sporophytes.

42. Combining ability of banana triploid hybrid progenitors and genomic prediction for agro-morphological traits

L. Toniutti$^{2,3,\alpha}$, *S. Rio*$^{1,2,\alpha}$, *C. Madec*^{2,3}, *S. Ricci*^{1,2}, *C. Guiougou*^{2,3}, *F. Marius*^{2,3}, *C. Mina*^{2,3}, *J.M. Delos*^{2,3}, *F. Lambert*^{2,3}, *J.C. Efile*^{2,3}, *A. D'Hont*^{1,2}, *G. Martin*$^{2,3,\beta}$, *J.Y. Hoarau*$^{2,4,\beta}$, *F. Salmon*$^{2,3,\beta}$

[1] CIRAD, UMR AGAP Institut, F-34398 Montpellier, France
[2] UMR AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro, F-34398 Montpellier, France
[3] CIRAD, UMR AGAP Institut, F-97130 Capesterre-Belle-Eau, Guadeloupe, France
[4] CIRAD, UMR AGAP Institut, F-97494 Sainte-Clotilde, La Réunion, France

In order to sustain the production of bananas, which is under threat from major diseases, it is essential to breed cultivars that are disease-resistant and also meet commercial standards. Since edible bananas are the final products of breeding and are seedless triploid hybrids, improving and combining parent plants is a crucial step. Currently, little is known about parental combining abilities and genomic prediction models that could inform banana breeding strategies.

In this study, a breeding population of 2,723 triploid individuals, derived from multiparental 2x-4x crosses over three crop cycles, was characterized. Twenty-three traits related to plant and fruit architecture and bunch yield were assessed. Phenotypic variance was partitioned into non-genetic and genetic effects, including the general combining ability (GCA) of diploid and tetraploid parents, the specific combining ability (SCA) and the genetic variability within crosses. The use of parental genomic information enabled cross mean performance prediction through genomic relationship matrices of general and specific combining abilities, the latter being partitioned into dominance and across-population epistasis contributions. Predictive abilities often greater than 0.5 were obtained, particularly when the tetraploid parent was observed in other crosses and, for some traits, when neither parent was observed. Information on genomic prediction of cross mean performance will help selecting and combining parents. A genomic prediction study using hybrid genotyping is ongoing on this dataset and will help streamline phenotyping efforts by targeting the most promising hybrids.

43. Combining ecophysiological and genetic modeling to identify new breeding targets for water stress tolerance in tomato

Fabien Tirado, *Gilles Vercambre*, *Emilie Millet*

Tomato is a mainly irrigated vegetable crop vulnerable to water shortage . With the increasing frequency of extreme climate events, incorporating drought tolerance in tomato breeding programs has become imperative. However, classical breeding targets, yield and fruit quality, are influenced by a multitude of factors affecting the plants throughout the growing season. Studying their response to stressing conditions remains challenging as it differs between genotype. So far, classical approaches either dissect the genetic architecture of agronomical traits measured on genetic diversity under stress, or study the physiological functioning of a unique variety under different conditions. We propose to combine disciplines by studying physiological traits derived from an ecophysiological model and their response to drought. We hypothesize that these traits are under simple genetic control so we can dissect their genetic architecture and identify candidate genes, making them new breeding targets for drought tolerance. We used an integrated ecophysiological tomato model to identify measurable traits enabling the calculation of these physiological traits. We phenotyped accordingly a genetically diverse core collection of 150 cherry tomato accessions in a semi-equipped greenhouse under control and drought conditions. We identified physiological traits with genetic variability and their response to drought. Combining with the whole genome re-sequencing information (6M SNP) we will dissect the genetic architecture of these traits using GWAS approaches and genomic prediction models. In the longer term, this integrative approach will allow development of new tomato ideotypes with enhanced tolerance to water stresses.

44. First application of genomic prediction in quinoa

Clara Stanschewski, *Mark Warmington*, *Irfan Afzal*, *Evan Craine*, *Kevin Murphy*, *Mark Tester*, *Jesse Poland*

Chenopodium quinoa is a semi-domesticated crop of growing global interest for its high nutritional value and natural tolerance to various adverse environmental conditions. For expansion into novel environments of cultivation, agriculturally important traits need to be assessed across different environments and targeted for selection in breeding programs. Genomic prediction is a tool that allows the prediction of the performance of untested genotypes, and the selection of best performing genotypes for target environments. To evaluate the applicability of genomic prediction in quinoa, a machine learning approach and a statistical approach were compared in a panel of 559 quinoa accessions for several agronomically important traits including plant height, days to flowering, thousand grain weight, and seed yield from three years of field trials in two locations, Australia and Pakistan. Prediction accuracies varied substantially by trait and location. Seed traits (TGW, seed length) predicted well in Australia ($r = 0.46-0.63$) but poorly in Pakistan ($r = 0.15-0.16$). Days to flowering showed better prediction in Pakistan using machine learning (ML) ($r = 0.52$ vs 0.33 GBLUP), while GBLUP performed better for Australia. Yield remained challenging, with modest correlations in Australia ($r = 0.34$ GBLUP, 0.24 ML) and near-zero in Pakistan ($r = 0.01-0.10$). While best performing method varied by trait, the machine learning approach delivered more consistent predictions and lower sensitivity to training population composition. These results provide a promising foundation for implementing genomic prediction to inform and accelerate quinoa breeding programs

45. Neodomestication of a halophytic perennial grain (*Distichlis palmeri*)

Izamar Olivas Orduna, Francisco Molina Freaner, Kashif Nawaz, Jesse Poland

Freshwater scarcity and soil salinization are impacting agriculture globally. Major crops exposed to mild salinity show yield losses of 50–80%. In contrast, halophytes are well-adapted to grow in full seawater. Given these challenges, the Neodomestication of new halophytic species offers a disruptive shift for using saline resources in agroecosystems.

Distichlis palmeri is a halophyte from the Poaceae family, stands out as a strong candidate for this purpose. It already produces a grain similar in size and nutritional value to rice and wheat. Native to the saltmarshes of the Gulf of California in Mexico, *D. palmeri* has long been harvested and consumed by indigenous people from the region.

We have collected germplasm of *D. palmeri* and used this plant material for genetic studies. Our dataset includes genotypic and phenotypic information from 600 individual plants from 14 populations. From these collections, we assembled high-quality reference genomes and observed differentiation of the populations across the Gulf of California. We delved into genetic association mapping for important agronomic traits such as plant height, spikelet/floret number, and spike size. In parallel, we are studying plant growth and developing cultivation systems that mimic natural saline environments, where we test different salinity levels. Within these systems, we estimated moderate to high heritability for traits like plant height, biomass, and greenness (NDVI).

The vision of this project is the neodomestication of *D. palmeri* as a halophytic grain crop. By establishing solid genomic and agronomic foundations we aim of converting this wild species into a cultivated saline crop.

46. Genome-wide insights into wheat flour properties and rheology: Genetic markers and candidate genes for improving end-use quality

Juan Menor de Gaspar[1] *, *Richard D. Cuthbert*[2], *Yuefeng Ruan*[2] *, *Ron Knox*[2], *Bin Xiao Fu*[3], *Kun Wang*[3], *Jatinder S. Sangah*[2], *Samia Berraies*[2], *Brad Meyer*[2], *Firdissa E. Bokore*[2], *Julio Isidro y Sánchez*[1] *

[1] *Centro de Biotecnologia y Genómica de Plantas (CBGP, UPM-INIA), Universidad Politécnica de Madrid (UPM) - Instituto Nacional de Investigación y Tecnologia Agraria y Alimentaria (INIA), Campus de Montegancedo-UPM, Pozuelo de Alarcón, 28223, Madrid, Spain.
[2] Agriculture and Agri-Food Canada, Swift Current Research and Development Centre, 1 Airport Road, Swift Current, SK S9H 3X2, Saskatchewan, Canada.
[3] Aquatic and Crop Resource Development, National Research Council of Canada, 110 Gymnasium Place, Saskatoon, SK S7N 0W9, Saskatchewan, Canada.

Improving the quality traits of bread wheat (*Triticum aestivum*) is crucial for meeting global food demands while maintaining end-use quality. This study utilizes genome-wide association studies (GWAS) to identify marker-trait associations (MTAs) influencing key quality traits, including protein content, milling yield, dough mixing energy, resistance to stretching, water absorption, and extensibility. Using genotyping-by-sequencing (GBS) data, 1,767 high-quality single nucleotide polymorphisms (SNPs) were analyzed across a diverse Canada Western Red Spring (CWRS) wheat panel, evaluated for different quality traits near Swift Current, Saskatchewan, Canada, from 2009 to 2019. 31 distinct MTAs were detected on 13 chromosomes, including known loci like Rht-B1 and Glu-1, and many novel associations highlighting wheat quality traits' polygenic nature. Additionally, we assessed the predictive performance of genomic selection (GS) models, comparing GBLUP and BayesB approaches, with and without incorporating significant GWAS hits as fixed effects. While the inclusion of GWAS hits resulted in minimal improvements in predictive accuracy, slight gains for water absorption suggest that this trait is governed by a fewer number of larger effect loci compared to the rest of the tested quality traits. Functional annotation of candidate genes near significant MTAs revealed links to stress response, protein metabolism, and grain-filling processes, providing insights into the genetic architecture of these traits. This study enhances our understanding of wheat quality genetics and identifies novel loci that present opportunities for targeted genetic improvement, while also demonstrating the potential of integrating GWAS and GS for breeding programs.

47. Quantifying protein levels in cereal–legume intercrops using near-infrared spectroscopy

Reena Dubey[1], *Anne Zaar*[2], *Malgorzata Niewinska*[3], *Riccardo Zustovi*[1], *Kevin Dewitte*[1], *Sofie Landschoot*[1], *Greet Verlinden*[1], *Geert Haesaert*[1], *Steven Maenhout*[1] *

[1]Department of Plants and Crops, Faculty of Bioscience Engineering, Ghent University, Ghent, Belgium
[2]Institute for Breeding Research on Agricultural Crops, Julius Kühn-Institute, Sanitz, Germany
[3]DANKO, Hodowla Roślin Sp. z o.o., Choryń 27, 64-000 Kościan, Poland

A cereal-legume intercrop represents a climate-resilient, low-input cropping system that enables a sustainable increase in protein yield per unit area of land. Intercropping triticale (× *Triticosecale*) and faba bean (*Vicia faba* L.) offers synergistic benefits such as improved nitrogen use efficiency, yield stability, and increased protein levels. However, high-throughput methods for evaluating crude protein levels, such as Near-Infrared Spectroscopy (NIRS), assume single-species seed or flour samples, requiring an explicit post-harvest separation of the produce. This study aims to explore the use of NIRS as a high-throughput method to quantify crude protein content and anti-nutritional compounds in triticale–faba bean mixtures.

To address the complexity of mixed samples, powdered grain mixtures were prepared in eleven defined proportions by weight, ranging from pure triticale to pure faba bean in decreasing (triticale) and increasing (faba bean) steps of 10%. Ten genetically diverse genotypes of each species were selected to reflect the trait variability that is encountered in a breeding pool. For each triticale and faba bean accession, protein content, as well as vicine and convicine concentrations—key anti-nutritional compounds in faba bean—were determined by means of wet chemistry methods. NIRS calibration models were optimised and validated for exceedingly larger subsets of mixing ratios, enabling an explicit quantification of the trade-off between generalizability and accuracy of these models.

Results indicate that a series of hierarchical NIRS calibration curves enable rapid and reliable estimation of protein and other key quality traits in triticale–faba bean mixtures. The presented approach is of interest to intercrop breeders as well as food and feed processors that rely on an on-site assessment of the nutritional composition of incoming raw materials.

48. Genetic evaluation and simulation of grain yield in Australia Southern faba bean breeding program

Huanhuan Zhao[1,2], *Samuel Catt*[1], *Charity Chidzanga*[1], *Julian Taylor*[1,2]

[1] School of Agriculture, Food and Wine, The University of Adelaide, Adelaide, SA, Australia
[2] Analytics for the Australian Grains Industry (AAGI), The University of Adelaide, Adelaide, SA, Australia

Understanding trends of genetic improvement within a breeding program is essential for enhancing breeding efficiency and integrating new breeding strategies. In this study, we estimate the genetic trends of grain yield within the National Faba Beans (*Vicia faba*) Southern Breeding Program using historical data collected from 2018 to of 2024. A total of 1985 entries and 158 field trials conducted at the 25 locations were analysed. The range of grain yield spanned from 0.033 t/ha to 11.200 t/ha with a mean of 2.738 t/ha. Grain yield exhibited a positive genetic trend of 0.145 t/ha across the breeding advancement stages (S1–S4) over the 2018–2022 breeding cycles. We further estimated the breeding values for the elite breeding lines (S4) by a factorial analytical model with the relationship matrix based on pedigree. The estimated genetic gain for the grain yield is around 0.1% per annum in the breeding program. Simulation studies comparing phenotypic selection based on pedigree with genomic selection (GS) at early and late stages demonstrate the advantages of GS, including reduced costs, higher selection accuracy within the product development pipeline, and the potential to significantly shorten breeding cycles in population improvement pipelines. Therefore, incorporating new technologies such as GS is essential to improve breeding efficiency and achieve higher genetic gains in faba bean breeding.

49. Dissecting the genetic basis of cold tolerance in potato through QTL mapping and functional analysis of CBF1

Beiyu Tu[1], *Li Shi*[1], *Rashmi Sasidharan*[2], *Richard G.F. Visser*[1], *Sara Bergonzi*[1]

[1]Plant Breeding, Wageningen University & Research, P. O. BOX 386, Wageningen, 6700 AJ, The Netherlands
[2]Plant Stress Resilience, Institute of Environmental Biology, Utrecht University, 3584 CH, Utrecht, TheNetherlands

Cold stress in early spring limits the productivity and establishment of potato, a major field crop. Although late planting can mitigate this stress, early planting with enhanced cold tolerance holds significant potential advantages. To investigate the genetic basis of cold tolerance, we performed QTL mapping in a diploid potato population grown in pots under natural cold conditions. The major QTL identified co-localized with a genomic region containing six copies of CBF (C-repeat binding factor) genes, known regulators of cold response. Allele mining in this region revealed substantial genetic variation, and we subsequently focused on CBF1 due to its strong allelic diversity and expression differences across the population. Sequence variation was identified in both the promoter and coding regions of CBF1. Expression QTL (eQTL) analysis indicated that CBF1 is under cis- regulation, suggesting that promoter variation likely drives expression differences. To further explore the functional impact of coding variation, we are currently comparing different CBF1 alleles through overexpression analysis. These results provide insights into the molecular regulation of cold tolerance in potato and may support marker-assisted breeding for improved stress resilience.

50. flexFitR and exploreHTP: Open-source software to enable nonlinear modeling of plant growth from remotely-sensed imagery

Johan Steven Aparicio, Jeffrey Endelman

In this study, we introduce two complementary R packages: exploreHTP, a Shiny-based application for extracting plot-level data from remotely sensed imagery, and flexFitR, a package for flexible nonlinear least squares model fitting. Their application is demonstrated through analysis of ground cover and plant height data from a potato breeding trial at the University of Wisconsin-Madison. RGB images were collected through eight uncrewed aerial surveys during the 2022 growing season, with exploreHTP enabling efficient extraction of phenotypic measurements. Four nonlinear models (logistic, linear-plateau, linear-logistic, and quadratic-plateau) were fitted to ground cover data using flexFitR. Based on the Akaike Information Criterion (AIC) and other penalized measures of fit, the quadratic-plateau model was optimal for 75% of the plots. Plant height was accurately modeled using a rapid-growth and exponential-decay structure. The area under the growth curves (AUC) for ground cover and plant height had similar predictive ability for tuber yield: $r = 0.46$ under cross-validation. When the two traits were combined in one linear model, the predictive ability for yield increased to $r = 0.56$ ($p < 0.05$). Penalized regression using the full time series offered no advantage over AUC-based predictions. In summary, the flexFitR and exploreHTP packages offer a practical way to characterize plant growth and support decision-making in breeding programs.

51. Haplotype stacking to improve stability of stripe rust resistance in wheat

Jingyang Tong (presenting author), Zerihun T. Tarekegn, Dilani Jambuthenne, Hannah Robinson, Madhav Pandit, Kira Villiers, Sambasivam Periyannan, Lee Hickey, Eric Dinglasan, Ben Hayes
Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, St Lucia 4072, Queensland, Australia

We investigated stripe/yellow rust (YR) responses in the Vavilov wheat diversity panel evaluated across 11 field experiments conducted in Australia and Ethiopia during 2014–2021. Genotype-by-environment interaction (GEI) was analysed using a factor analytic (FA) model. Genotype-level selection was performed with overall performance (OP) and root-mean-square deviation (RMSD), which reflected average performance and stability of YR resistance across environments, respectively. Genomic estimated breeding values (GEBV) for these traits were calculated and compared with those from a multi-trait GBLUP model with average performance represented by the mean GEBV across environments and stability by the standard deviation of GEBV across environments. The FA-based and multi-trait GBLUP GEBV had high correlations. Haplotypes with large effects on OP and RMSD were identified using the local GEBV method. Favourable haplotypes were then used for stacking in breeding simulations, using the Vavilov collection as a base. Compared to truncation selection, optimal haplotype selection (OHS) using an artificial intelligence (AI)-based algorithm achieved longer-term genetic gains for both OP and RMSD (after many generations) by initially selecting founder parents that maximised favourable haplotypes. Simulations using YR responses from diverse environments that mimicked fluctuating environmental conditions across seasons were conducted to evaluate strategies for selection of YR resistance that is stable across years. Strategies which gave most weight to OP, but some weight to RMSD were optimal in these conditions, and substantially reduced variation of performance across years. This study provides useful information for breeding cultivars with both high YR resistance and high stability of resistance across environments.

52. Harnessing plant genetic resources without compromising local adaptation in hybrid and inbred wheat breeding programs

Maegan Green[1,3,4], *Gregor Gorjanc*[2], *Ian Henderson*[1], *Tally Wright*[3], *Carus John-Bejai*[4]

[1]Department of Plant Sciences, University of Cambridge, Cambridge, United Kingdom
[2]The Roslin Institute and Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Easter Bush, Midlothian, United Kingdom
[3]Niab, Park Farm, Histon, Cambridge, United Kingdom
[4]Wheat Breeding, KWS UK Ltd, Thriplow, United Kingdom

Global food security is increasingly threatened by climate change and increasing population-driven demand, requiring the enhancement of genetic diversity in wheat breeding programs. Plant genetic resources (PGRs), including wild relatives, landraces, and isolated breeding gene pools, represent a largely underutilised host of diversity with potential to improve yield and resilience. However, the introgression of beneficial traits from PGRs into elite germplasm remains constrained by cost and complexity. This study develops and evaluates simulation-based strategies combined with general combining ability (GCA) estimation to identify elite parental lines that facilitate the effective introgression of PGR-derived diversity without compromising local adaptation. Emphasis is placed on key adaptive traits, plant height, flowering time, and growth habit, controlled by major loci Rht, Ppd1, and Vrn1. A tester-based crossing design is employed to generate training populations that enhance GCA prediction accuracy and facilitate diversity capture in both inbred and hybrid breeding contexts. Simulation outputs will be validated against empirical data to assess predictive performance. Additionally, a decision-support R-based tool is under development to optimise diversity capture in backcross breeding schemes, enabling efficient incorporation of PGRs while preserving local adaptation. This integrative framework offers a resource for enriching genetic diversity in wheat breeding pipelines through the strategic use of simulation and predictive breeding tools.

53. Dissecting genotype × environment interactions for cold tolerance traits in sorghum using a haplotype-based framework

Mohamed Mosalam[1,2]*, *Hannah Robinson*[1], *Rod Snowdon*[2], *Kai Voss-Fels*[1]

[1]Department of Plant Breeding, Hochschule Geisenheim University, Geisenheim, Germany
[2]Department of Plant Breeding, Justus Liebig University Giessen, Giessen, Germany

Understanding genotype × environment interactions is crucial for enhancing sorghum resilience in temperate climates. In this ongoing study, we employ a genome-wide haplotype framework to dissect the genetic architecture of cold tolerance by integrating multi-environment phenotypic data. Haplotype blocks were first delineated within the five sorghum botanical races (Caudatum, Kafir, Durra, Guinea, and Bicolor) and then projected onto a recombinant inbred line (RIL) population derived from inter-racial crosses among these races. Cold tolerance was assessed across three developmental stages: juvenile, seedling, and reproductive. Statistical analyses were conducted in ASReml-R, fitting multi-environment mixed models for juvenile and reproductive traits and a repeated-measures model for seedling vigour to partition additive and G×E variance.

We identified both stable and environment-specific haplotypes associated with cold responses, with several key regions tracing back to specific founder races. This approach enhances the resolution of G×E analyses and has the potential to support genomic selection strategies for breeding cold-resilient sorghum cultivars adapted to temperate environments.

54. Leveraging a broad gradient of plant-plant interactions to efficiently breed for cereale-legume mixtures

Jemay Salomon, *Jérôme Enjalbert*, *Timothée Flutre*

Background: Crops are increasingly exposed to abiotic and biotic stresses. Intercropping—growing multiple species together—offers improved yield stability, enhanced disease control, and, particularly for cereal-legume mixtures, reduced reliance on nitrogen fertilizers in low-input systems. Due to the required redesign of the farming system and strict market standards, farmers prefer sole cropping. Consequently, plant breeders have largely overlooked intercropping. Moreover, plant-plant interactions can affect mixture yield, highlighting the need to develop adapted varieties. However, the combinatorial complexity of assembling N×N binary mixtures from N genotypes per species remains a major challenge for breeding. We hypothesized that jointly analyzing sole- and intercropping data from incomplete designs would improve the estimation of variance components and genetic values.

Methods: We adapted a genetic model of varietal mixtures for intercropping, integrating a "within species interaction (SIS)" and developing a statistical model. We then conducted simulations with realistic variance components to assess (co)variance estimates and selection reliability by exploring : i) genetic parameters, especially, the trade-off between direct breeding values (DBV)-social breeding values (SBV) and weight of SIS, ii) complete versus sparse designs with varying space allocated to mixtures. We also established two similar field trials—one fully and partially replicated—to evaluate the proposed framework and compare simulation outcomes with real data.

Results and Conclusion: Our findings from simulation show that incomplete experimental designs can accurately estimate genetic values, and genotype selection efficiency varies with SIS variance relative to DBV variance. The method will be demonstrated using results from an ongoing field experiment on wheat-pea intercropping.

55. Harnessing heterosis in faba bean: Breeding strategy optimisation through stochastic simulations

Amanda Karlström[1], *Diana Bengtsson*[2], *Andreas Hansson*[2], *Alf Ceplitis*[2], *Aakash Chawade*[1], *Åsa Grimberg*[1]

[1]Swedish University of Agricultural Sciences (SLU), Sundsvägen 5, 234 56 Alnarp
[2]Lantmännen, Lundqvists väg 11, 268 31 Svalöv

Faba bean (*Vicia faba* L.) is a high-protein legume well-suited to cool climates; however, breeding progress is constrained by its partially outcrossing mating system, susceptibility to inbreeding depression, and the absence of a stable hybrid seed production system. Synthetic variety development has been proposed as a promising approach to harness heterosis in this species, though most empirical studies have focused on short-term genetic gains. In this study, we employed stochastic simulations using AlphaSimR to compare long-term genetic gains from two breeding strategies: pedigree-based line breeding and synthetic variety development. We also evaluated the impact of incorporating genomic selection (GS) into each strategy. To investigate the influence of heterosis, simulations were conducted across three levels of dominance, representing varying degrees of non-additive genetic effects. Results highlight the potential of synthetic variety development to exploit heterosis and show that GS can enhance genetic gain, particularly under moderate to high dominance scenarios. This work provides a framework for optimizing faba bean breeding strategies, accounting for species-specific challenges and genetic architecture.

56. Enhancing QTL detection and genomic prediction for stay-green in maize using UAV-based multi-trait analysis

Yan-Cheng Lin[1], *Hany Elsharawy*[2], *Kang Yu*[2], *Claude Urbany*[3], *Antonina Shlykova*[3], *Armin Hölker*[3], *Pascal Schopp*[3], *Milena Ouzunova*[3], *Thomas Presterl*[3], *Sebastian Urzinger*[1], *Manfred Mayer*[1], *Chris-Carolin Schön*[1]

[1]Plant Breeding, TUM School of Life Sciences, Technical University of Munich, 85354 Freising, Germany
[2]Precision Agriculture, TUM School of Life Sciences, Technical University of Munich, 85354 Freising, Germany
[3]Maize Breeding, KWS SAAT SE & Co. KGaA, 37574 Einbeck, Germany

Multi-trait (MT) models improve the identification of pleiotropic or closely linked loci for QTL detection. In genomic prediction, they enhance accuracy by leveraging genetic correlations among traits. High-throughput phenotyping technologies, such as unmanned aerial vehicles (UAVs), generate large volumes of data that can be integrated into MT models to improve the analysis of complex traits. In this study, we evaluated the line per se performance of 388 maize doubled haploids (DH) lines from an eight-founder MAGIC population across four environments, two locations over two years. One of the sites had sandy soil and rainfed management to induce drought stress. Visual stay-green scores were recorded at 2, 4, and 6 weeks after flowering, while UAV imagery was collected weekly from flowering until the final stay-green assessment. The DH lines were genotyped using whole-genome shallow sequencing, yielding 356,668 SNPs. We investigated the construction of various stay-green-related vegetation indices, as well as strategies for selecting UAV-derived traits and applying multi-trait models. Several indices exhibited higher heritability than the visual stay-green scores. Additionally, multi-trait GWAS revealed novel QTL that were not detected in single-trait analyses. Incorporating UAV-derived traits as auxiliary traits in genomic prediction models improved the accuracy of predicting stay-green performance in untested genotypes.

57. Discovering optimal genotype–environment combinations in rice through a Bayesian optimization-like Randomr forests

Hideto Mochizuki (Grad. Sch. Agr. Life Sci., Univ.),
Kosuke Hamazaki (Adv. Int. Proj., RIKEN),
Chikashi Sato (Ifu-Rinrin),
Akira Abe (Iwate Biotechnology Research Center),
Chyon Hae Kim (Iwate University, Sky Ocean Technology Co., Ltd. Sky Ocean Technology Co., Ltd.),
Hiroyuki Shimono (Iwate University),
Hiroyoshi Iwata (Grad. Sch. Agr. Life Sci., Univ.)

Developing climate-resilient crop varieties is essential for stable agricultural production under future climate change. While genomic selection (GS) accelerates breeding programs, it often fails to predict phenotypes accurately in untested environments because of genotype-by-environment interactions (G×E). Bayesian optimization (BO) has been proposed as a promising approach to account for G×E by modeling prediction uncertainty. However, traditional BO using Gaussian processes becomes computationally expensive with high-dimensional genomic and environmental data. In this study, we propose a novel BO-inspired method utilizing Random Forests (RF), where the variance of predictions across decision trees is used as a proxy for uncertainty. We applied this method to a large-scale dataset of rice cultivars (238 genotypes) tested over 37 years across 103 locations in Japan. In each environment, we iteratively selected 10 genotype–environment combinations based on four acquisition functions: predicted mean ($\hat{y}$), probability of improvement (PI), expected improvement (EI), and upper confidence bound (UCB). This process was repeated for 20 cycles. We also employed k-medoids clustering to ensure diversity among the selected genotypes. Results showed that RF with clustering achieved performance comparable to BO (reaching optimal solutions in $\sim$14 cycles using PI, EI, or UCB), while reducing computation time by over 50%. RF models remained stable ($\sim$1.5 seconds per cycle) regardless of iteration count, in contrast to BO, which scaled up to 21 seconds per cycle. Our findings demonstrate that RF-based methods offer a practical, scalable alternative to conventional BO for G×E-aware crop breeding.

58. Identification of maize heterotic group-specific haplotypes and impact of residual inbreeding on elite hybrids grain yield

Romain Kadoumi[1,2], *Nicolas Heslot*[1], *Fabienne Henriot*[1], *Alain Murigneux*[1], *Mathilde Berton*[1], *Laurence Moreau*[2], *Alain Charcosset*[2]

[1] Limagrain Field Seeds, 28 Route d'Ennezat, 63720 Chappes, France
[2] Génétique Quantitative et Evolution – Le Moulon, Institut National de la Recherche Agronomique et Environnementale, Université Paris-Sud, AgroParisTech, Université Paris-Saclay, F-91190 Gif-sur-Yvette, France

Modern hybrid maize (*Zea mays* L.) breeding programs revolve around the management of distinct complementary heterotic groups to maximize heterosis in high-performing hybrids. This practice lowers shared genetic backgrounds and maximizes divergence to limit inbreeding in hybrids. However, most breeding programs have not always enforced strict separation between heterotic groups in the past.

This study proposes a new haplotype-based approach to evaluate hybrids' residual inbreeding based on parental similarity. This method identifies a negative effect on hybrids' grain yield more significant than raw SNP data. Analysis of modern experimental hybrids revealed an inbred family contributing to superior rates of residual inbreeding. Assessment of these inbreds revealed haplotype transfers between heterotic groups, originating from the use of a Stiff Stalk-Iodent commercial hybrid as source breeding material.

Competitor commercial hybrids are a distinctive source of genetic variation for inbred development, which can inherently affect crossing pools divergence. In this study, the introduction of this uncharacterized intergroup parent increased heterotic group-specific haplotype migration between crossing pools. These fragments caused significant genome-wide residual inbreeding in experimental hybrids across selection cycles.

This study highlights the necessity for accurate evaluation of introduced breeding start materials to minimize haplotype migration and admixture between crossing pools. We demonstrate the consequences of using commercial hybrids in inbred development, particularly regarding haplotype sharing, residual inbreeding, and their effects on hybrid performance. Insights from these results can assist breeders in optimizing inbred breeding start crosses progenitors in a reciprocal recurrent selection scheme while introducing genetic diversity.

59. Joint analysis of monovarietal and mixed stand performances to study the genetic architecture of indirect genetic effects in wheat cultivar mixtures.

M. Remérand, J. Enjalbert, T. Flutre

Université Paris-Saclay, INRAE, CNRS, AgroParisTech, GQE, Le Moulon, Gif-sur-Yvette 91190, France

Varietal mixtures represent an important agronomical tool, notably to increase yields under lower pesticide use and provide other ecosystemic services. They also contribute to stabilize yield thanks to compensation but other biological mechanisms can play a role, such as complementarity and facilitation. However, the underlying genetic architecture of such ecological interactions in mixtures is still lacking.

We designed a two-season, alternate rank field trial to investigate the Direct Genetic Effects (DGE) and Indirect Genetic Effects (IGE) of a hundred wheat recombinant lines from a MAGIC population. We analysed jointly pure stands and 3-components varietal mixtures on multiple traits.

We found evidence for a small yet positive overyielding in mixtures, concurring with current literature, further suggesting varietal mixtures are an interesting agronomical lever to increase yields.

Linear mixed modeling was used to separate DGEs and IGEs from the non-heritable residual components. A GWAS was then performed on the estimated DGEs and IGEs using a 420k-SNP microarray. Several SNPs were associated with DGEs for grain yield and height, and one was associated with IGEs on grain weight. This preliminary result prompts more investigation on the role of indirect effects in mixtures, but is promising to start incorporating such effects in selection schemes for mixing ability.

60. EasyGeSe – a resource for benchmarking genomic prediction methods

Daniel Ariza-Suarez, *Carles Quesada-Traver*, *Bruno Studer*, *Steven A. Yates*

Molecular Plant Breeding, Institute of Agricultural Sciences, ETH Zurich, Zurich, Switzerland

Genomic prediction is a widely used method for estimating breeding values from genotypic data and uses a range of models for prediction. Although the performance of new models is usually evaluated based on common measurements of predictive accuracy, these assessments are usually tested on individual datasets, limiting generalizability. To address this, we present EasyGeSe, a resource that allows for the standardized evaluation of genomic prediction methods using a curated collection of diverse real-world datasets. The datasets cover 93 traits from ten plant and animal species, including barley, common bean, lentil, loblolly pine, eastern oyster, maize, pig, rice, soybean and wheat. These data are provided in analysis-ready formats and can be loaded using accompanying functions from dedicated R and Python packages, which enable their integration into existing workflows. We benchmarked this dataset using a range of ten commonly used models, including parametric, semi-parametric and non-parametric approaches, and provide the corresponding 5×5-fold cross-validation splits and prediction results. This setup enables direct and fair comparisons between new methods and established baselines under consistent conditions. By harmonizing input data and evaluation procedures, EasyGeSe promotes reproducibility and transparency in performance assessments of genomic prediction models. Its coverage of multiple species and traits ensures broad applicability, helping users select appropriate models for their specific contexts. It also serves as a practical teaching tool, offering accessible datasets with paired genotypic and phenotypic information for hands-on learning.

61. Can we teach machines to select like a plant breeder? A recommender system approach to support early generation selection decisions based on breeders’ preferences

Sebastian Michel, Franziska Löschenberger, Christian Ametz, Herbert Bistrich Hermann Bürstmayr

Plant breeding is considered to be the science and art of genetically improving plants according to human needs. Breeders in this context oftentimes face the difficult task of selecting among thousands of genotypes for dozens of traits simultaneously. Using a breeder’s selection decisions from a commercial wheat breeding program as a case study, this study investigated the possibility of implementing a recommender system based on the breeder’s preferences to support early-generation selection decisions in plant breeding. The target trait was the retrospective binary classification of selected versus non-selected breeding lines during a period of five years, while the selection decisions of the breeder were predicted by various machine learning models. The explained variance of these selection decisions was of moderate, and the models’ precision suggested that the breeder’s selection decisions were to some extent predictable, especially when some of the pending selection candidates were part of the training population. Training machine learning algorithms with breeders’ selection decisions can thus aid breeders in their decision-making processes, particularly when integrating human and artificial intelligence in the form a recommender system to potentially reduce a breeder’s effort and the required time to find interesting selection candidates.

62. Quantifying the drivers of genetic change in plant breeding

T.P. Oliveira, D.J. Tolhurst, B. Poupard, G. Gorjanc

Quantifying how germplasm inflow shapes genetic progress is essential for sustainable breeding. We extend pedigree-based partitioning to resolve the additive mean and variance contributed by imported lines while explicitly modelling genotype-by-environment interaction (GEI). Using AlphaSimR, we simulated two parallel hybrid-maize programmes, BP1 and BP2, each evaluated across seven stages. BP2 received the top sixteen lines from BP1 each year for ten years (short import), forty years (long import) or not at all, under scenarios of no, low and high GEI. A Bayesian linear mixed model estimated stage means; an LDL decomposition of the pedigree matrix attributed genetic change to local and imported pathways.

Under no or low GEI, a single decade of import elevated the mean and variance of BP2 to levels comparable with the four-decade no-import baseline, accelerating long-term gain without exhausting diversity. High GEI nullified this benefit: the short import had negligible impact on the asymptotic mean, and only continuous inflow maintained elevated variance, with diminishing returns and a switch to positive covariance signalling maladaptation. Concordance correlations between true and estimated partitions exceeded 0.7 across scenarios, confirming methodological robustness.

The resulting framework provides breeders with a cost-effective tool to forecast the genetic consequences of germplasm exchange, align import duration with GEI intensity, and balance gain against diversity. Although demonstrated in maize, the approach is crop-agnostic and applicable to any multi-environment breeding programme where strategic introgression under environmental complexity is pivotal.

63. CrossingTools: an R package for mating optimization in plant breeding

Sven E. Weber[1], *Daniel J. Tolhurst*[2], *Dominic L. Waters*[2], *Christian R. Werner*[3,4]

[1]Justus-Liebig University Giessen, Giessen, Germany
[2]The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian, UK
[3]Accelerated Breeding Initiative (ABI), Consultative Group of International Agricultural Research (CGIAR), Texcoco, Mexico
[4]International Maize and Wheat Improvement Center (CIMMYT), Texcoco, Mexico

Mate allocation is a critical component of plant breeding programs, which aims to generate superior progenies and improved populations over time. While numerous methodologies for mate allocation exist, few software tools offer a unified and scalable framework that integrates these key methodologies in a user-friendly manner. CrossingTools addresses this gap by providing a flexible, efficient solution for optimizing mating decisions across a wide range of breeding schemes. The package supports genomic BLUP and selection indices for a wide range of variance structures and implements a variety of selection criteria to evaluate the potential of a cross, including the expected mean, optimal haploid value and the superior progeny value. For the latter, different methods to calculate the additive and dominance segregation variances for both inbred and heterozygote parents are implemented, which makes CrossingTools well-suited to any breeding scheme. After criteria calculation, the user can rank all potential crosses based on the selected criterion or apply CrossingTools' optimal crossing selection routine, which balances genetic gain and genetic diversity to support long-term improvement. The package is implemented within the R statistical environment using C++ subroutines, allowing a user-friendly scripting language while ensuring high computational performance and scalability for large populations with dense marker data. The flexibility and utility of CrossingTools are demonstrated through simulation studies comparing different selection criteria in both classical line breeding programs and two-part breeding schemes.

64. AI-based predictions in canola

Iulian Gabur[1,2], *Tiberiu Sarbu*[1], *Danut Petru Simionuc*[1], *Rod Snowdon*[2]

[1] "Ion Ionescu de la Brad" Iasi University of Life Sciences, Department of Plant Breeding, Iasi, Romania
[2] Justus Liebig University, Department of Plant Breeding, Giessen, Germany

In recent years, the introduction of modern varieties and genetic advancements have made plant breeding one of the most dynamic agricultural sectors. Progress has been further accelerated by the recent introduction of big-data analytics and artificial intelligence (AI), which have improved yields and adaptability in a variety of environments. Predictive breeding is the foundation of AI-based selection, which makes use of sizable, intricate datasets from modern programs. However, individual breeders trying to train reliable AI models face significant challenges due to the high-dimensionality and heterogeneity of breeding data, such as that from canola programs. New analytical pipelines and machine-learning techniques that can handle genomic, phenomic, and environmental inputs concurrently, are needed to address these complexities. Feature selection and overfitting control are essential for implementing AI successfully. Techniques such as PCA and Random Forest–based feature selection (RF-FS) are essential for lowering dimensionality and enhancing model generalisability. AI models enhanced by these methods routinely outperform conventional methods in canola breeding. Overall, AI-driven tools are revolutionizing plant breeding by integrating complex datasets, reducing computation time, and enhancing predictive accuracy. These advancements empower breeders to make more informed decisions, leveraging genotypic data more effectively to develop resilient and high-performing varieties.

65. Identification of QTL involved in heavy metal and trace element accumulation in durum wheat

Alexia Crézé

Arvalis – Institut du vegetal

Wheat is a major staple crop and a key source of both essential micronutrients and toxic heavy metals in the human diet. Among these, cadmium (Cd) stands out due to its toxicity and the greater propensity of durum wheat to accumulate it in the grain. In response to growing health concerns, the European Union recently lowered the regulatory limits for Cd in durum wheat. To support the development of cultivars with reduced heavy metal accumulation while maintaining adequate levels of beneficial elements in durum wheat, we quantified the grain concentrations of 15 elements. This study investigated the genetic basis of heavy metal and trace element accumulation in wheat using genome-wide association studies (GWAS). A panel of 136 durum wheat cultivars was grown in field trials in France in 2019. Grain concentrations of toxic elements (As, Cd, Cr, Ni, Pb) and essential nutrients (Ca, Co, Cu, Fe, Mg, Mn, Mo, P, K, Zn) were measured. While some elements were below detection thresholds, phenotypic data for detectable elements enabled the identification of 9 QTLs associated with their accumulation, spread across chromosomes 1B, 2A, 3B, 4B, 5A, 5B, and 7B. Several of these QTLs colocalized with previously reported loci or candidate genes involved in metal transport and homeostasis. Bioinformatic analyses further explored gene annotations and the functional relevance of associated regions. These findings provide valuable insights for breeding wheat varieties with lower heavy metal content and improved nutritional quality, in line with increasingly strict food safety regulations.

66. Insufficient memory: a practical problem associated with the genomic relationship matrix and estimation of GxE effects for grain yield across multiple years and sites

Wallace Cowling[1], *Renu Saradadevi*[1], *Felipe Castro-Urrea*[1,2], *Clare Mukankusi*[3]

[1] The University of Western Australia, Perth, Australia
[2] NPZ Australia Pty Ltd, Perth, Australia
[3] Alliance of Bioversity International and International Center for Tropical Agriculture (CIAT), Kampala, Uganda

In our research on rapid recurrent genomic selection in common bean (Saradadevi et al. 2021), we aimed to develop genomic analysis models for grain yield (GY) that were useful in a regional breeding program for common bean in East Africa, in a project funded by the Australian Centre for International Agricultural Research. Our goal was to develop an animal-model style of crop breeding, where data for GY and other traits accumulated over years, and approximately 1,000 new genotypes were added each year with low concurrency of genotypes across trial sites and years. Sparse testing for GY occurred in 33 spatially-designed field trials in 6 countries over 4 years. Factor analytic (FA) models were implemented to estimate the impact of genotype x environment (GxE) interactions on GY in a one-stage analysis with genomic and/or pedigree relationship information and optimised spatial models at each site.

One-stage FA genomic models were often too complex, and reasonably-equipped laptop computers had insufficient memory to solve the models. Interestingly, if the genomic relationship matrix was replaced with a pedigree relationship matrix, we were able to complete the one-stage FA analysis and investigate the impact of GxE on breeding values. This confirmed that the “insufficient memory” problem was specifically associated with the use of the genomic relationship matrix.

We seek a practical solution which allows breeders with limited computing resources to exploit pedigree and genomic relationship information and optimise spatial models in each environment, explore GxE and maximise genetic gain through single step genomic analysis.

Reference: *Saradadevi et al (2021) Multivariate genomic analysis and optimal contribution selection predicts high genetic gains in cooking time, iron, zinc and grain yield in common beans in East Africa. Plant Genome 14:e20156. https://doi.org/10.1002/tpg2.20156*

67. Integration of environmental predictors and genomic selection into a pipeline for global potato variety recommendation

Marco Peixoto[1], *Liza Van der Laan*[1], *Lenin Pereira*[2], *Tiago Marçal*[2], *Leo Hoffmann*[1], *Lincoln Zotarelli*[3], *Henrique Boriolo*[4], *Gerrit Hoogenboom*[4], *Kaio Dias*[5], *Marcio Resende*[1]

[1]Sweet corn and Potato Breeding and Genomics lab, University of Florida, Gainesville, Florida, USA
[2]Departamento de Biologia, Universidade Federal de Lavras, Lavras, Minas Gerais, Brazil
[3]Horticultural Science Department, University of Florida, Gainesville, Florida, USA
[4]Global Food Systems Institute & Agricultural & Biological Engineering, University of Florida, Gainesville, Florida, USA
[5]Departamento de Biologia Geral, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil

The integration of environmental, genetic, and statistical tools offers a powerful approach for optimizing cultivar recommendations under increasing environmental uncertainty. Here, we present a pipeline for optimizing cultivar selection and recommendations of potato cultivars on a global scale and across future scenarios. Yield (YT) and specific gravity (SG) traits were collected from 110 environments (spanning 2011–2023 years and 11 locations across the United States), which included over 800 genotypes. Genotypic data included 15K markers used to construct an additive relationship matrix. Environmental predictors (EP) were derived from soil and climate datasets covering the 110 training locations and new environments (NeEnv) on a global scale. Over 400 EPs were retrieved from raw data, contributing to refined environmental modeling. The NeEnv were accounted for in a 100 km² matrix worldwide, yielding 61000 environments. We implemented and compared, within the training set, different models with genomic and EPs, and genotype-by-interaction information, such as factor analytic and reaction norms, in both single- and two-stage analyses. Model performance was assessed using leave-one-environment-out cross-validation. In the second step, we combined the trained model and EPs to generate a worldwide prediction for the NeEnv. The optimal planting window for each NeEnv was predicted using crop growth models, and it was used to generate EPs accordingly. A dissimilarity matrix was built, spatial interpolation was used to obtain a contiguous representation, and the trained model was used to predict genotype performance in the NeEnv. The correlations for the best model had an average of 0.55 (YT trait – 0.17 to 0.70) and 0.70 (SG trait – 0.24 to 0.89). Maps were generated to highlight top-performing cultivars, overall performance trends, similarity of NeEnv to training environments, and the projected impact of environmental changes on cultivar selection. The pipeline implemented here is therefore an effective method for cultivar recommendation.

68. Breed-E-Omics European project: Genomics and genetics approaches of Spelt (Triticum spelta) for a sustainable agriculture facing global warming

Eduardo D. Munaiz[*,1], *Julio Isidro y Sànchez*[2], *François Buche3*, *Sameh Selim*[1], *UniLaSalle*[4], *Lemaire-Deffontaines*[5], *Chambre d'agriculture de l'Oise*[6], *Moulins Waast7*, *Céréales Experts*[8], *Nicolas Gallois*[3], *Céline Roisin*[1], *Julien Guidet*[1]

[1]UniLaSalle, Collège Agrosciences, AGHYLE Research Unit, 60000 Beauvais, France
[2]Centre for Plant Biotechnology and Genomics UPM, Campus de Montegancedo. Madrid, Spain
[3]UniLaSalle, Collège Agroalimentaire et Bioprocédés, IDEALIS Research Unit, Beauvais, France
[4]UniLaSalle, 19 Rue Pierre Waguet, Beauvais, France
[5]Lemaire-Deffontaines SAS, 59310 Auchy-lez-orchies, France
[6]Chambre d'agriculture de l'OISE, Rue Frere Gagne, Beauvais, France
[7]Moulins Waast SARL, 400 rue du moulin, Mons-en-pevele, France
[8]Céréales Experts, 24 Rue d'Anjou, Saulgé l'Hopital, 49320 Brissac Loire Aubance, France

The European project 'Breed-epeautre-omics' aims to study the genetics and quality improvement of spelt (*Triticum spelta*, in French epeautre) in a context of ecological transition and global warming as a source of alternative crops for producers in France (Hauts-de-France) and Spain. A context in which, for instance, 2022 was the hottest year in the history of France, and crop's yields (e.g. wheat, maize, etc.) were considerably reduced strongly impacting the income and competitiveness of producers. In 'Breed-epeautre-omics' we propose an innovative solution to this challenge integrating genetics, quantitative genetics, transcriptomic studies, genomic prediction, and high-throughput qualitative analysis. 'Breed-epeautre-omics' is a collaborative initiative that integrates all stakeholders of the cereal sector composed of public and private research institutions, regional public agencies, cereal transformation stakeholders, seed breeding companies, and growers.

Acknowledgements:
We gratefully acknowledge the support of the European Innovation Partnership grant, PEI Haute-de-France DADR-2024-029390, within the European FEADER program and the co-financing of the Region Haute-de-France.

List of contributors

A. D’Hont, 119
A. Reza Sharifi, 90
Aakash Chawade, 132
Aaron Beattie, 22
Adam Serghini, 83
Aiyesa Leke Victor, 40
Akira Abe, 134
Akito Kaga, 104
Alain Baranger, 42
Alain Charcosset, 23, 34, 93, 135
Alain Murigneux, 135
Albrecht E. Melchinger, 73
Alessandro Vannozzi, 116
Alex Lenkoski, 106
Alexander Heinlein, 61
Alexandre Colmant, 82
Alexia Crézé, 142
Alexis Comar, 83
Alf Ceplitis, 132
Alice Boizet, 103
Allan Cummiskey, 22
Alvarez-Jarreta, J., 102
Amanda Karlström, 132
Amine Abbadi, 85
Andrea Doeschl-Wilson, 46
Andrea Patocchi, 29
Andreas Hansson, 132
Andreas Stahl, 85
Andrew Olson, 70
Angela S. Prudencio, 63
Annaïg De Walsche, 34
Anne Zaar, 124
Anne-Katrin Gorn, 86
Antero Heikkilä, 35
Antonina Shlykova, 133
Antonio Augusto Franco Garcia, 108, 117
Antonio R. Jiménez, 63
Armin Hölker, 73, 133
Arun Isaac, 112
Asa Grimberg, 132
Ashild Ergon, 118
Asuka Itaya, 22
Atikur Rahman, 105
Augustin Desprez, 23

B. Gardunia, 87
B. Poupard, 139
Baoxing Song, 70
Barry Harrison, 114
Bedford, J.A., 102
Beiyu Tu, 126
Ben Hayes, 74, 128
Benjamin Eppler, 47
Benjamin Pommerrenig, 85
Bernadette Julier, 42
Bernard De Baets, 69
Betina Lado, 107
Bin Xiao Fu, 123
Bjarne Nielsen, 30
Boyny Zsa Zsa, 100
Brad DeHaan, 22
Brad Meyer, 123
Brezo Mateos, 89, 101
Brian Gardunia, 49
Bright Enogieru Osatohanmwen, 90
Bruna Marques Moreno, 108
Bruno Bernazeau, 42
Bruno Desprez, 23
Bruno Santos, 110
Bruno Studer, 68, 137

C.M. Richardson, 87
C. Adams, 87
C. Guiougou, 119
C. Madec, 119
C. Mina, 119
C. Robin Buell, 115
Carel Peeters, 55
Carles Quesada-Traver, 68, 137
Carlos Robles Zazueta, 39, 84
Carlos Viquez-Zamora, 103
Carolina Rivera-Poulsen, 73
Carus John-Bejai, 129
Cathal McCabe, 105
Chambre d’agriculture de l’Oise, 145
Charity Chidzanga, 125
Charlene Wight, 22

Chikashi Sato, 134
Chris Gaynor, 49
Chris-Carolin Schön, 73, 133
Christian Ametz, 138
Christian Flachenecker, 85
Christian R. Werner, 20, 21, 91, 140
Christina Finegan, 49
Christina Lehermeier, 49
Christine Te Riini, 110
Christine Tollon-Cordet, 42
Chyon Hae Kim, 134
Clara Polzer, 73
Clara Stanschewski, 121
Clara Voiry, 42
Clare Mukankusi, 143
Claude Urbany, 133
Claudia Serrano Colomé, 113
Clemens Flamm, 56
Clément Bienvenu, 64
Cockram, J., 102
Connell, J., 102
Corné Verburg, 61
Cristina Martínez-Andújar, 63
Céline Roisin, 145

Dagne Wegary, 20, 91
Dan MacEachern, 22
Daniel Ariza-Suarez, 137
Daniel Tolhurst, 21, 49, 53, 111, 139, 140
Daniel Valle Torres, 85
Daniela Bustos-Korts, 54, 56
Danut Petru Simionuc, 141
David Gonzalez Dieguez, 20, 91
David Moody, 78
David Pot, 64
Dean Dellapenna, 115
Denis Pageau, 22
Di Wu, 115
Diana Bengtsson, 132
Diana Carolina Solarte Certuche, 117
Diana Marcela Pérez-Valencia, 60
Dilani Jambuthenne, 128
Dini Ganesalingam, 78
Dirk Reheul, 69
Do Than Dat Le, 88
Dominic Waters, 49, 53, 111, 140
Dorcus C. Gemenet, 91
Doreen Ware, 70
Duncan Henderson, 111
Dyer, S., 102

Eder Jorge de Oliveira, 117
Eduardo D. Munaiz, 145
Eleni Nikolaidou, 70
Elisa De Luca, 116
Elsa Gonçalves, 75
Emanuele De Paoli, 116
Emi Tanaka, 27
Emilia Koch, 47
Emilie Millet, 120
Eric Dinglasan, 128
Espen Sørensen, 105
Evan Craine, 121

F. Lambert, 119
F. Marius, 119
F. Salmon, 119
Fabiano Pita, 82
Fabien Tirado, 120
Fabienne Henriot, 135
Falola, O., 102
Felipe Castro-Urrea, 143
Finn Gaida, 113
Fiona Doohan, 105
Firdissa E. Bokore, 123
Flint, B., 102
Florence Parat, 56
Francesco Flagiello, 116
Francisco Molina Freaner, 122
Francisco Pérez-Alfocea, 63
Franck Gauthier, 34
Franz Goecke, 118
Franz-Leopold Haupt, 85
Franziska Löschenberger, 138
François Buche3, Sameh Selim, 145
Fred van Eeuwijk, 50, 54, 55, 60, 61, 80

G. Martin, 119
Gabriel Mamedio de Freitas, 117
Gard W. Gravdal, 99
Gary Atlin, 20, 91
Geert Haesaert, 124
Geneviève Telmosse, 22
Gennady Bracho-Mujica, 56
George Terry Epaku, 114
Gerrit Hoogenboom, 144
Gilles Vercambre, 120
Giorgetti, S., 102
Giovanny Covarrubias-Pazaran, 82
Glenn Bryan, 92
Greet Verlinden, 124

Gregor Gorjanc, 49, 129, 139
Gregory R. Schoenbaum, 115
Guilherme da Silva Pereira, 108
Guillaume P. Ramstein, 70
Guillermo Sniadower, 107
Gunnar Klemetsdal, 118
Guoliang Li, 33, 81

Hamish A. Burnett, 99
Hanna Haikka, 105
Hannah Robinson, 39, 78, 84, 128, 130
Hans-Jürgen Auinger, 73
Hans-Peter Piepho, 28, 39, 47, 56, 78, 86, 94
Hany Elsharawy, 133
Hao Cheng, 38, 112
Harrison, P., 102
Helen Booker, 22
Henrik Jensen, 99
Henrique Boriolo, 144
Herbert Bistrich, 138
Hermann Bürstmayr, 138
Hideto Mochizuki, 134
Hilde Halland, 105
Hiroyoshi Iwata, 104, 134
Hiroyuki Shimono, 134
Hisashi Tsujimoto, 104
Holly Byker, 22
Hrannar Smari Hilmarsson, 105
Huanhuan Zhao, 125
Hugo Dorado, 50
Hugo Gangloff, 88
Hugues de Verdal, 64
Huihui Li, 79, 96
Hélène Muranty, 29

Ian Henderson, 129
Indalécio Cunha Vieira Júnior, 90
Inés Rebollo, 107
Irfan Afzal, 121
Isabelle Morasse, 22
Ismo Strandén, 35
Iulian Gabur, 141
Izamar Olivas Orduna, 122
Iñigo Azqueta, 113

J.C. Efile, 119
J.M. Delos, 119
J.Y. Hoarau, 119
J. Enjalbert, 41, 136
J. Salomon, 41
Jacques David, 42
Jagroop Kahlon, 24
Jakob Streuber, 85
James Cockram, 67
James Walter, 62
Janek Sven-Ole Sendowski, 70
Jatinder S. Sangah, 123
Javier Fernandez, 54
Javier Fernández-González, 36, 76
Jean-Luc Jannink, 117
Jeffrey Endelman, 127
Jemay Salomon, 131
Jenn Walker, 24
Jennifer Mitchel-Fetch, 22
Jens Hartung, 47, 86
Jesse Hemerik, 54
Jesse Poland, 121, 122
Jianming Yu, 115
Jihad Orabi, 30
Jill Cairns, 20
Jingyang Tong, 128
Jip Ramakers, 54
Jochen C. Reif, 33, 81
Johan Steven Aparicio, 127
John Crowley, 24
John Moore, 110
John P. Hamilton, 115
Joost van Heerwaarden, 50
Jorge Cadima, 75
Jose Ángel Martín-Rodríguez, 63
Joseph O. Ogutu, 94
Joshua C. Wood, 115
Juan Burgeno, 20
Juan José Guerrero, 63
Juan Menor de Gaspar, 123
Jude Sise, 110
Judith Frigeau-Reid, 22
Juho Hautsalo, 105
Julian Ramirez-Villegas, 50
Julian Taylor, 62, 66, 125
Julie Aubert, 88
Julie Graham, 89, 101
Julien Frouin, 103
Julien Guidet, 145
Julio Isidro y Sànchez, 145
Julio Isidro y Sánchez, 36, 76, 123
Junita Solin, 80, 83
Justin Blancon, 48

Jérémy Labrosse, 83
Jérôme Bartholomé, 93
Jérôme Enjalbert, 131
Jørgen Ødegård, 118

K.R. Grant, 87
Kai Ilves, 105
Kai Voss-Fels, 39, 74, 78, 84, 130
Kaio Dias, 144
Kali Stewart, 22
Kang Yu, 133
Karen McLean, 92
Karen Wolf, 28
Karine Henry, 23
Kashif Nawaz, 122
Katharine F. Preedy, 89
Keith Gardner, 72
Kelly Houston, 114
Kengo Sakurai, 104
Kevin Dewitte, 124
Kevin Murphy, 121
Killian Melsen, 55
Kira Villiers, 74, 128
Klaus Norhausen, 35
Kosuke Hamazaki, 104, 134
Kristy Hobson, 53
Kun Wang, 123

L. Busswinkel, 87
L. Toniutti, 119
Laura E. Tibbs-Cortes, 115
Laura Zanin, 116
Laurence Moreau, 23, 83, 93, 135
Leanne Fischbuch, 24
Lee Hickey, 53, 78, 128
Lemaire-Deffontaines, 145
Lenin Pereira, 144
Leo Hoffmann, 144
Leopoldo Sanchez, 93
Lester Nicholas, 100
Leticia A. de C. Lara, 24, 87
Li Shi, 126
Lidija Bitz, 105
Lincoln Zotarelli, 144
Liza Van der Laan, 144
Lodha, D., 102
Love, B., 102
Lucia Gutierrez, 48, 105, 107
Luigi Falginella, 116
Luke Ramsay, 67

Líllian Beatriz Januario Bibiano, 108

M. Post, 87
M. Remérand, 41, 136
Madhav Pandit, 128
Mads Sønderkær, 92
Maegan Green, 129
Maialen Ormazabal, 63
Mainassarah Zaman-Allah, 20
Malcolm Macaulay, 67
Malgorzata Niewinska, 124
Manfred Mayer, 73, 133
Manon Verdier, 101
Manuel Derrien, 109
Marc Labadie, 83
Marcio Resende, 144
Marco Peixoto, 144
Margaret Wallace, 67
Maria Lucia Carneiro Vieira, 108
Maria Magallanes-Lundback, 115
Maria Thorkildsen, 106
Marion Buffard, 103
Marjorie Hanneman, 115
Mark Tester, 121
Mark Warmington, 121
Martin Boer, 54, 60, 80
Martin Howes, 24
Martin Lidauer, 35
Martin Quincke, 107
Maryse Hulmel, 109
María Xosé Rodríguez-Álvarez, 60
Mason Chizk, 24, 110
Massel Karen, 100
Mathieu Tiret, 93
Mathilde Berton, 135
Matthew Hayes, 22
Maud Tenaillon, 23
Maximilian Schmidt, 39
Mehri-Hadinezhad, 22
Melinda Drummond, 22
Michael A. Gore, 115
Michaela Jung, 29, 68
Michelle Teran-Pineda, 73
Mike Holzworth, 22
Mikhail Beloglasov, 42
Milena Ouzunova, 73, 133
Milka Malenica, 85
Millicent Smith, 53
Min Lin, 105, 106
Minoru Inamori, 104

Mohamed Mosalam, 130
Mohd Ibnur Syawal Zakaria, 97
Morten Lillemo, 105, 106
Moulins Waast7, Céréales Experts, 145

Naamati, G., 102
Natalie Howes, 24
Natasha H. Johansen, 70
Nathalie Boissot, 34
Nathan Mountain, 22
Nazanin Zamani-Noor, 85
Neil Budko, 61
Nicholas T. Deason, 115
Nicholas Tinker, 22
Nicola Rossi, 98
Nicolas Gallois, 145
Nicolas Heslot, 135
Nicolas Salas, 42
Nolan Bornowski, 115
Nora Røhnebæk Aasen, 106
Nuri Güvencli, 84

Oliver Bitz, 105

P. Amer, 87
Pablo González Barrios, 107
Pablo Sandro, 48, 107
Pascal Schopp, 32, 133
Paul Shaw, 67
Paula Silva, 107
Percival-Alwyn, L., 102
Pernilla Vallenback, 105
Pernille Bjarup, 30
Pernille Sarup, 30
Peter Crisp, 39
Peter Skov Kristensen, 30
Philip Greenspoon, 101
Pierre Devaux, 23
Pierre Fernique, 28
Powell Owen, 100
Purificación A. Martínez-Melgarejo, 63

R. Bouchon, 87
Rachele Falchi, 116
Rajiv Sharma, 98
Rashmi Sasidharan, 126
Reena Dubey, 124
Renate H. Schmidt, 33, 81
Renaud Rincent, 83, 88
Renu Saradadevi, 143
Rex Bernardo, 95
Riccardo Zustovi, 124
Richard D. Cuthbert, 123
Richard G.F. Visser, 126
Richard Horsnell, 67
Richard Mott, 112
Rishap Dhakal, 107
Robert D. Hancock, 89
Rod Snowdon, 130, 141
Romain Kadoumi, 135
Ron Knox, 123
Rosemary Bailey, 26
Ryokei Tanaka, 115

S. Larmer, 87
S. Ricci, 119
S. Rio, 119
Salvador Gezan, 54, 55
Sambasivam Periyannan, 128
Samia Berraies, 123
Samir Alahmad, 78
Samuel Catt, 125
Sanjeev Kumar Sharma, 92
Sara Bergonzi, 126
Sara Taskinen, 35
Saraf, S., 102
Sarah Ghysels, 69
Sarah Hearne, 72
Savka Orozovic, 22
Savvas Chatzivasileiou, 70
Scott Chapman, 54
Sebastian Michel, 138
Sebastian Urzinger, 133
Sebastian Warnemünde, 85
Seifelden M. Metwally, 76
Shang Gao, 96
Shanice Van Haeften, 53
Sheh May Tam, 97
Shirin Mohammadi, 106
Shuai Wang, 70
Signe Bråtelund, 118
Sigridur Dalmannsdottir, 105
Silius Mortensønn Vandeskog, 106
Silvina Baraibar, 78
Simon Rio, 103
Smaragda Tsairidou, 111
Snowdon Rod, 100
Sofia Pereira da Silva, 73
Sofie Landschoot, 124
Stefan Scholten, 40

Stefanie Muff, 99
Stephan Hartmann, 86
Stephanie Brunner, 78
Steven Maenhout, 69, 124
Steven Penfield, 52
Steven Yates, 68, 137
Stéphanie Sidibe-Bocs, 103
Susan McCallum, 101
Sven Weber, 85, 140
Syafeqa Abdul Hamid, 97
Séverine Monnot, 109
Sónia Surgy, 75

T. Byrne, 87
T. Flutre, 41, 136
T. Oliveira, 87, 139
Talissa Oliveira Floriani Zimermman Souza, 108
Tally Wright, 67, 102, 129
Tansy Chia, 67
Teshale Mamo, 20
The participants of the CGIAR Allele Mining project, 72
Therese Birkeland Fossøy, 105
Thomas Bataillon, 70
Thomas Kreuter, 85
Thomas Presterl, 73, 133
Tiago Marçal, 144
Tiberiu Sarbu, 141
Tim Beissinger, 40
Timo Strack, 39
Timothy Beissinger, 90
Timothy George, 114
Timothy Thavarajah, 62
Timothée Flutre, 131
Tom Kirk, 24
Tom Ruttink, 118
Tom Theeuwen, 80
Tristan Mary-Huard, 34, 88

UniLaSalle, 145

Valeria De Rosa, 116
Valeria Montano, 101
Vanda M. Lourenço, 94
Vanessa McMillan, 67
Vincent Garin, 103
Vincent Segura, 64
Vivi Arief, 54

W. Bourdoncle, 87
Wallace Cowling, 143
Walter Chivasa, 20
Wayne Powell, 98
Weber Sven, 100
Weikai Yan, 22
Wera Eckhoff, 56
Wolfgang Link, 40
Wubishet Bekele, 22

Xabi Cazenave, 93
Xiaowei Li, 115

Yan-Cheng Lin, 133
Yong Jiang, 33, 81
Yoseph Beyene, 20, 91
Yue Lin, 110
Yuefeng Ruan, 123
Yusheng Zhao, 33
Yusuke Toda, 104

Zachary Aldiss, 78
Zerihun T. Tarekegn, 128