AGBT2024™

AGRICULTURE

Element Biosciences

AVITI™ Transforming Agricultural Genomics

Breeding and genome engineering programs need genotyping solutions that move fast and cultivate strong returns. Join our Gold Sponsor talk, stop by our booth, and visit us at the Element Suite while at AGBT-Ag. We're looking forward to discussing your low-pass sequencing and genotyping-by-sequencing needs and how the low cost and simple analysis of the Element AVITI System can help. Get ready to Shift Your Science.

Gold Sponsor Talk

Marine Conservation in the Age of Cost-effective Genomics: The Leatherback Turtle

Monday, April 15, 2024
12:00 – 12:20pm
Komatke

Booth #201
Element suite: Gila Monster

AGBT 2024™ AGRICULTURE

We are honored to welcome you to the Advances in Genome Biology and Technology (AGBT) 3rd annual Ag meeting. You are among the world's leading genome researchers, data scientists, breeders, policy influencers, funders, and technology innovators from the global North and South embracing opportunities to redesign terrestrial and aquatic agriculture.

Over the next few days, we will focus on the integration of genomics and agriculture to address the escalating needs of a changing Earth.

AGBT Ag is the third global meeting organized by The Genome Partnership, a Saint Louis based non-profit that advances research, promotes education and expands commerce in genome science and technology. The Genome Partnership has been proudly organizing AGBT meetings since 1999.

Organizing Committee

We are incredibly grateful to our meeting Co-Chairs Sarah Hearne, Daniela Lourenco and the entire Scientific Organizing Committee for the countless hours they invested in making AGBT Ag24 a content-rich experience.

Over the next several days we will gather to collaborate, share the latest scientific discoveries, and discuss new technological breakthroughs thanks in large part to the work of our organizing committee.

Xiaofeng Cao
Principal Investigator-Institute of Genetics and Developmental Biology, Chinese Academy of Sciences (CAS)

Jack Dekkers
Distinguished professor, section leader of animal breeding and genetics, Iowa State University

Appolinaire Djikeng
Director General, International Livestock Research Institute (ILRI), Senior Director, CGIAR Livestock-based systems

Sarah Hearne
Genetic Resources Program Director a.i. (GPR), CIMMYT

Charlie Johnson
Executive Director Genomics & Bioinformatics, Texas A&M AgriLife

Renee Lafitte
Deputy Director for Crops R&D, Bill and Melinda Gates Foundation

Nathan Lakey
President and Chief Executive Officer, Orion Genomics

Daniela Lourenco, University of Georgia
Associate Professor, Animal and Dairy Science, University of Georgia

Len Pennacchio
DOE Joint Genome Institute, Lawrence Berkeley National Laboratory

Alison Van Eenennaam
Extension Specialist, Animal Biotechnology and Genomics, University of California, Davis

Bruce Walsh
Professor of Ecology and Evolutionary Biology, University of Arizona

Susan Wessler
Distinguished Professor of Genetics, University of California, Riverside (Ret.)

AGBT2024™
AGRICULTURE

We gratefully acknowledge the generous support provided by the following companies

GOLD

Element Biosciences

SILVER

Complete GENOMICS™

EXHIBITORS AND CONTRIBUTING SPONSORS

AgriPlex GENOMICS
ALITHEA GENOMICS
ARIZONA Core Network
BECKMAN COULTER Life Sciences
biomodal
FERRIS GENOMICS
HUDSONALPHA INSTITUTE FOR BIOTECHNOLOGY
mirxes
NEOGEN
NEW ENGLAND Biolabs
Oxford NANOPORE Technologies
PacBio
PARAGON GENOMICS
PHASE GENOMICS
SGS
ThermoFisher SCIENTIFIC
The world leader in serving science

General Information

Registration / Hospitality Desk

Sunday, April 14	7:00 a.m. – 7:00 p.m.
Monday, April 15	7:30 a.m. – 7:00 p.m.
Tuesday, April 16	7:30 a.m. – 7:00 p.m.
Wednesday, April 17	7:30 a.m. – 7:00 p.m.

Name Badges

Please wear your name badge & lanyard at all times. Security will refuse entry to anyone not associated with the AGBT meeting. The sharing of a name badge will result in you being asked to leave the conference.

Transportation

Phoenix Sky Harbor International Airport (PHX) is 15 miles from the Sheraton Grand at Wild Horse Pass. The transfer time to the resort is approximately 20 minutes, depending on traffic. Shuttles will depart every 30 minutes at the top of the hour and half-past the hour during the following times:

Arrivals:	Sunday, April 14	1:00 p.m. – 8:00 p.m.
Departures:	Thursday, April 18	4:00 a.m. – 2:00 p.m.

Social Events

Welcome Dinner – **Monday, April 15 – 6:30 p.m. – 9:30 p.m.** – Akimel Lawn.

Sponsor Social – **Tuesday, April 16 – beginning at 7:35 p.m.** – Akimel Foyer. Grand Prize will be awarded during Wednesday's Plenary Session. (This prize is non-transferable.)

Farewell Dinner Party – **Wednesday, April 17 – 7:15 p.m. – 10:00 p.m.** - Hemapik Lawn.

Conference Technology

Download our official AGBT conference mobile app to access the agenda and abstracts easily. Check your email for the link to download the app, or visit our registration desk for assistance.

Start Posting!

#AGBTAg

Complete GENOMICS™

Complete Sequencing Solution For Every Lab

Sample & Library Prep

Lab Automation

DNBSEQ™ Sequencers

Software & Data Analysis

DNBSEQ™ Technology

- ✓ No Clonal Error
- ✓ No Index Hopping
- ✓ Low Duplication Rate

E25

2.5-7.5 Gb

G99

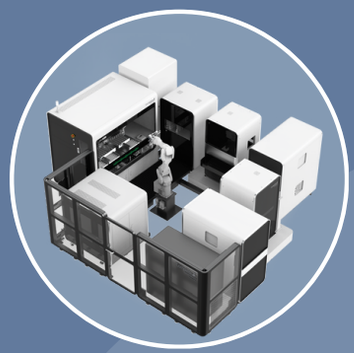
8-48 Gb

G400

55-1,440 Gb

T7

1,000-7,000 Gb



T20x2

8,000-72,000 Gb

Throughput (Gb/run)

10 50 1500 7,000 80,000

Agenda

Sunday, April 14	**Transfers 1 p.m. – 8 p.m. Phoenix Sky Harbor International Airport (PHX)**
7:00 a.m. – 7:00 p.m.	**AGBT Meeting Registration** Akimel Foyer
4:00 p.m. – 6:00 p.m.	**The Future of Farm Animal Genomic Research** Sergey Koren, NIH, Fiona McCarthy, Darren Hagen, Juan Steibel and Angelica Van Goor Akimel Ballroom 4
7:00 p.m. – 9:00 p.m.	**Getting Over the T2T Finish Line** Sergey Koren, Associate Investigator, NIH/NHGR Genome Informatics Section Arang Rhie, Staff Scientist, NIH/NHGR Genome Informatics Section Akimel Ballroom 4
Monday, April 15	**Scientific Program**
6:30 a.m. – 7:30 a.m.	**Rise and Shine with AGBT – Optional Activity** Meet at Hospitality Desk
7:30 a.m. – 7:00 p.m.	**Meeting Registration / Hospitality Desk** Akimel Foyer
7:30 a.m. – 9:00 a.m.	**Breakfast** Akimel Patio & Mesquite Terrace
9:00 a.m. – 9:10 a.m.	**Opening Remarks, (Sarah Hearne, Genetic Resources Program Director a.i. (GPR), CIMMYT and Co-Chair of the AGBT Ag Scientific Organizing Committee)** Komatke Ballroom
Plenary Session I:	**Global Opportunities and Challenges for Agriculture, I** (Sarah Hearne, Genetic Resources Program Director a.i. (GPR), CIMMYT and Co-Chair of the AGBT Ag Scientific Organizing Committee)

9:10 a.m. – 9:40 a.m.	**Ben Hayes, Queensland Alliance for Agriculture and Food Innovation, Centre Director, Animal Science, The University of Queensland** "Chromosome segment stacking to create ultimate crop and livestock genotypes"
9:40 a.m. – 10:10 a.m.	**Damaris Odeny, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Principal Scientist** “Exploiting climate resilience of dryland crops using advanced breeding tools”
10:10 a.m. – 10:30 a.m.	**Andre Garcia, (Abstract Selected) American Angus Association** “World Angus Evaluation, A collaboration to provide genetic evaluations for Angus cattle across the globe”
10:30 a.m. – 11:00 a.m.	**Coffee Break Exhibits** Akimel Ballroom 1-3
Plenary Session II:	**Global Opportunities and Challenges for Agriculture, II** (Nathan Lakey, Orion Genomics, Session Chair) Komatke Ballroom
11:00 a.m. – 11:30 a.m.	**Ross Houston, Benchmark Holdings, Director of Genetics and Innovation** “Aquaculture breeding: the next frontiers”
11:30 a.m. – 12:00 a.m.	**Howard Shapiro, Agricultural Sustainability Institute, UC Davis** "The speed of change is faster than our response: Nutrition & the end of stunting"
12:00 p.m. – 12:20 p.m.	**Gold Sponsor Talk - Element Biosciences** **John B. Horne, NRC Fellow, Senior Research Associate, South-west Fisheries Science Center, NMFS, NOAA** "Marine conservation in the age of cost-effective genomics: The leatherback turtle"
12:20 p.m. – 2:00 p.m.	**AGBT Lunch** Akimel Patio & Mesquite Terrace
12:40 p.m. – 2:10 p.m.	**Workshops and Lunch** Komatke Ballroom

12:40 p.m. – 2:10 p.m.	**Single Cell Technology in Plants and Animals Workshop and Lunch** **John Driver, Marc Libault and Xiaosa XU** Discuss updates to cell atlases being built in plants and animals, computational best practices or new ways to analyze this data, and examples of its use to test novel hypotheses.
2:10 p.m. – 2:55 p.m.	**Dessert & Coffee with Exhibits** Akimel Ballroom 1-3
Plenary Session III:	**Applying the Science, l** **(Jack Dekkers, Distinguished Professor, section leader of animal breeding and genetics, Iowa State University)** Komatke Ballroom
3:00 p.m. – 3:30 p.m.	**Diane Saunders, John Innes Centre, Group Leader, Head of Department** "Safeguarding wheat yields from cereal fungal invaders"
3:30 p.m. – 4:00 p.m.	**Samuel Oyola, Senior Scientist, Head of Genomics, International Livestock Research Institute (ILRI)** "Genomic epidemiology of rift valley fever virus"
4:00 p.m. – 4:20 p.m.	**Kathryn McRae, (Abstract Selected) AgResearch** “Pneumonia, Pseudopithomyces and Parasites: the role of genomics in addressing animal health challenges in sheep”
4:20 p.m. – 4:35 p.m.	**Silver Sponsor Talk - Complete Genomics** **Wenchi Liao, Director of Product Marketing, Complete Genomics** “Complete sequencing solutions for every lab” **Bryan Witherbee, Founder, Ferris Genomics** “Powering your breeding programs with MORE samples, MORE data” Komatke Ballroom
4:35 p.m. – 5:05 p.m.	**Coffee Break and Exhibits** Akimel Ballroom 1-3
Plenary Session IV:	**Precision Genetic Technologies** **(Alison Van Eenennaam, Extension Specialist, Animal Biotechnology and Genomics, University of California, Davis, Session Chair)** Komatke Ballroom

5:05 p.m. – 5:35 p.m.	**Ermias Kebreab, University of California, Davis, Associate Dean** "CRISPR and other innovations in reducing livestock methane emissions"
5:35 p.m. – 6:05 p.m.	**Erica McGale, University of Lausanne, UNIL** "Using genomic and epigenomic variation of both mycorrhizal fungi and cassava to improve yields in the tropics"
6:30 p.m. – 9:30 p.m.	**Welcome Dinner** Akimel Lawn
Tuesday, April 16	**(Morning)**
7:30 a.m. – 7:00 p.m.	**Hospitality/Information Desk** Akimel Foyer
7:30 a.m. – 9:00 a.m.	**Breakfast** Akimel Patio & Mesquite Terrace
Plenary Session V:	**Informatics and Data Analytics I (Daniela Lourenco, Associate Professor, Animal & Dairy Science, University of Georgia, Co-Chair of the AGBT Ag Scientific Organizing Committee, and Session Chair)** Komatke Ballroom
9:00 a.m. – 9:30 a.m.	**Charlie Messina, University of Florida** “Breeding for climate change is imperative but may occur differently as we think”
9:30 a.m. – 10:00 a.m.	**Eric Psota, PIC North America, Digital Innovation Manager** "Never stop improving: Advancing pig genetics using digital phenotyping"
10:00 a.m. – 10:20 a.m.	**Jenna Kalleberg, (Abstract Selected) University of Missouri** “Overcoming limitations to deep learning in domesticated animals with TrioTrain”

10:20 a.m. – 10:40 a.m.	**Gustavo de los Campos, (Abstract Selected) Michigan State University** "Regularization and transfer learning in genomic models through gradient descent with early stopping"
10:40 a.m. – 11:10 a.m.	**Coffee Break and Exhibits** Akimel Ballroom 1-3
Plenary Session VI:	**Informatics and Data Analytics II: (Bruce Walsh, Professor of Ecology and Evolutionary Biology, University of Arizona, Chair)** Komatke Ballroom
11:10 a.m. – 11:40 a.m.	**Mario Calus, Wageningen University, Associate Professor, Animal Breeding and Genomics** "Added value of functional annotation versus individual omics in genomic prediction"
11:40 a.m. – 12:00 p.m.	**AGBT Next Gen Leadership Awards**
12:00 p.m. – 2:00 p.m.	**AGBT Lunch** Akimel Patio & Mesquite Terrace
12:20 p.m. – 1:50 p.m.	**Data Sharing While Protecting Confidential and Proprietary Content Workshop and Lunch** **Hao Cheng, University of California, Davis** **Jack Dekkers, Iowa State University** **Angelica van Goor, USDA-NIFA** **Mario Calus, Wageningen University and Senior Editor of Genetics** **Radu Totir, Corteva Agriscience** **Rachel Hawken, Cobb-Vantress, Inc** The workshop will focus on current and new directives by governments and journals for sharing research data and how these can be aligned with the need to protect confidential or proprietary information that may be contained in the data. Akimel Ballroom 4
12:20 p.m. – 1:50 p.m.	**AgT2T - All Species Workshop and Lunch** Arang Rhie, Tom Goldammer, Tim Smith and Katie Jenike In this workshop you will hear from speakers about strategies and the benefits of telomere to telomere (T2T) genome assemblies for several agricultural species. Komatke Ballroom

1:50 p.m. – 2:55 p.m.	**Dessert & Coffee with Sponsors** Akimel Foyer
	Len Pennacchio, DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, Session Chair Komatke Ballroom
3:05 p.m. – 3:25 p.m.	**Julianna LeMieux, Deputy Editor-in-Chief, Genetic Engineering & Biotechnology News (GEN)** “Life after the bench: Adventures in science communication”
Plenary Session VII:	**From FASTQ to Fork (Len Pennacchio, DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, Session Chair)** Komatke Ballroom
3:25 p.m. – 3:55 p.m.	**Cathie Martin, John Innes Centre** "Healthy plants; healthy people - fortifying vitamins and phytonutrients in tomato"
3:55 p.m. – 4:25 p.m.	**Geoff Morris, Colorado State University** "From genome discovery to smallholder farmers - are we delivering?"
4:25 p.m. – 4:45 p.m.	**Elizabeth Ross, (Abstract Selected) The University of Queensland** “Opportunities and limitations of using long-read sequencing for predictive genomics”
4:45 p.m. – 5:05 p.m.	**Poster Flash Talks**
5:15 p.m. – 7:30 p.m.	**Posters & Palate Pleasers** Akimel Ballroom 1-3
Beginning at 7:35 p.m.	**Sponsor’s Social (Connect, Collaborate, and Chill)** Akimel Foyer
Wednesday, April 17	**(Morning)**
7:30 a.m. – 7:00 p.m.	**Hospitality/Information Desk** Akimel Foyer

7:30 a.m. – 9:00 a.m.	**Breakfast** Akimel Patio & Mesquite Terrace
Plenary Session VIII:	**Unlocking Innovation From Basic Science l, (Appolinaire Djikeng, Director General, International Livestock Research Institute (ILRI), Senior Director, CGIAR Livestock-based systems, Session Chair)** Komatke Ballroom
9:00 a.m. – 9:30 a.m.	**Elinor Karlsson, Broad Institute of MIT and Harvard, Director of the Vertebrate Genomics Group** "Investigating the validity of breed stereotypes through ancestry-inclusive genomics in thousands of dogs"
9:30 a.m. – 10:00 a.m.	**Giles Oldroyd, University of Cambridge, Director, Crop Science Centre** "Achieving sustainable productivity in agriculture through beneficial microbial associations"
10:00 a.m. – 10:20 a.m.	**Edward Symons, (Abstract Selected) Wellcome Sanger Institute** “Generating high-quality reference genomes assemblies at scale: The Tree of Life Genome Engine”
10:20 a.m. – 10:40 a.m.	**Lucas Rocha Moreira, (Abstract Selected) University of Massachusetts, Chan Medical School** “Inferring the genomic basis of cellular robustness in placental mammals”
10:40 a.m. – 11:10 a.m.	**Coffee Break and Exhibits** Akimel Ballroom 1-3
Plenary Session IX:	**Unlocking Innovation From Basic Science ll (Susan Wessler, Distinguished Professor of Genetics, University of California, Riverside, Session Chair)** Komatke Ballroom
11:10 a.m. – 11:40 p.m.	**Nathan Springer, Bayer, Head of Genetic and Genomic Innovation** "Opportunities to advance crop germplasm development using genome editing"

11:40 p.m. – 12:00 p.m.	**Keith Slotkin, (Abstract Selected) Donald Danforth Plant Science Center** "The control of transposable elements in plants"
12:00 p.m. – 12:20 p.m.	**Ricarda Jahnel, (Abstract Selected) University of Guelph** "Phenotyping for dairy cattle resilience: Enhancing livestock sustainability through genetics and genomics"
12:20 p.m. – 2:00 p.m.	**AGBT Lunch** Akimel Patio & Mesquite Terrace
12:40 p.m. – 2:10 p.m.	**Functional Annotation of Animal Genomes (FAANG) Workshop and Lunch (Chair, Dr. Martien Groenen, Wageningen University, The Netherlands)** Dr. Emily Clark, Roslin Institute, Edinburgh, UK Dr. Richard Crooijmans, Wageningen University, The Netherlands Dr. Lingzhao Fang, Arhus University, Denmark Dr. Irene Kaplow, Carnegie Mellon University, Pittsburgh, USA Dr. Martien Groenen, Wageningen Universtity FAANG celebrates its 10th anniversary to discover basic functional knowledge of genome function to decipher the genotype-to-phenotype (G2P) with an emphasis on farmed animals and its applications in animal breeding. It provides fundamental insight in molecular mechanisms underlying complex traits, adaptive evolution, domestication, and speciation and as such overlaps with other large ongoing projects like the Farm animal Genotype-Tissue Expression (FarmGTEx) project and the Zoonomia project aimed at discovering the genomic basis of shared and specialized traits in mammals. This workshop is aimed at exploring and strengthening further collaborations between these different ongoing projects.

Plenary Session X:	**Quantitative and Molecular Genetics (Charlie Johnson, Executive Director Genomics & Bioinformatics, Texas A&M AgriLife, Session Chair)** Komatke Ballroom
2:45 p.m. – 3:15 p.m.	**Mike Gore, Cornell University, Chair of the Plant Breeding and Genetics Section in the School of Integrative Plant Science** "A genome-level approach to balancing the micronutrient content of maize grain"
3:15 p.m. – 3:45 p.m.	**Xiaofeng Cao, Chinese Academy of Sciences** "Breeding of new forage grass for saline-alkali land improvement"
3:45 p.m. – 4:05 p.m.	**Jingyang Tong, (Abstract Selected) The University of Queensland** “Haplotype level characterization of resistance traits in Vavilov wheat accessions reveals the potential for breeding multi-disease-resistant varieties”
4:05 p.m. – 4:25 p.m.	**Owen Hudson, (Abstract Selected) University of Florida** “Developing genomic selection models, association studies, and machine vision algorithms to predict Fusarium ear rot severity”
4:25 p.m. – 4:55 p.m.	**Coffee Break** Komatke Foyer
Plenary Session XI:	**From the Cutting Edge-New Ideas and Opportunities (Renee Lafitte, Deputy Director for Crops R&D, Bill and Melinda Gates Foundation, Session Chair)** Komatke Ballroom
5:00 p.m. – 5:20 p.m.	**Tatiana Briody, (Abstract Selected) Queensland Alliance for Agriculture and Food Innovation, The University of Queensland** “High throughput precision editing at bovine genomic safe harbours”

5:20 p.m. – 5:40 p.m.	**Ivan Liachko, (Abstract Selected) Phase Genomics, Inc.** "Building the world's largest virome- and resistome- host interaction atlas for precision microbiomics"
5:40 p.m. – 6:00 p.m.	**Evan Ernst, (Abstract Selected) Howard Hughes Medical Institute; Cold Spring Harbor Laboratory** "The genomes and epigenomes of aquatic plants (Lemnaceae) promote triploid hybridization and clonal reproduction"
6:00 p.m. – 6:20 p.m.	**Temitayo Olagunju, (Abstract Selected) University of Idaho** "An ovine pangenome to characterize genetic diversity in sheep breeds"
6:20 p.m. – 6:40 p.m.	**Maci Mueller, (Abstract Selected) Kansas State University** "Long-read amplicon sequencing reveals large on-target deletions in CRISPR/Cas9 gene-edited bovine samples"
6:40 p.m. – 6:50 p.m.	**Closing Comments and Click2Win Announcements (Sarah Hearne, Principal Scientist, International Maize, and Wheat Improvement Center) (CIMMYT) and AGBT Ag Conference Co-Chair**
7:15 p.m. – 10:00 p.m.	**Farewell Dinner and Stargazing** Hemapik Lawn
Thursday, April 18	**Transfers 4 a.m. – 2 p.m. Phoenix Sky Harbor International Airport (PHX)**

AGBT2024™

PRECISION HEALTH

SEPTEMBER 4-6, 2024

THE GAYLORD ROCKIES, DENVER, COLORADO

SAVE THE DATE

LEARN MORE BY SCANNING THE QR CODE,
OR VISIT WWW.AGBT.ORG

Abstract Table of Contents

page

Quantitative and Molecular Genetics

A genome-wide association study of stress hormone levels in hair of young healthy pigs

Fazhir Kayondo[1], Hayder Al-Shanoon[2], Yolande M. Seddon[2], David Janz[2], Dylan Carette[2], Carmen Cole[2], Frederic Fortin[3], John C. S. Harding[2], Michael K. Dyck[4], Graham Plastow[4], Pig Gen Canada[5], Jack C. M. Dekkers[1]

1 Iowa State University, Department of Animal Science, Ames, IA
2 University of Saskatchewan, Western College of Veterinary Medicine, Saskatoon, Saskatchewan, Canada
3 Centre de Devéloppement du Porc du Québec Inc., Québec City, Quebec, Canada
4 University of Alberta, Department of Agricultural, Food and Nutritional Science, Edmonton, Alberta, Canada
5 Pig Gen Canada Research Consortium, Guelph, Ontario, Canada

The levels of stress hormones (cortisol, cortisone, DHEA, and DHEA-S) in the hair of young and healthy pigs are being evaluated for their potential as genetic indicators of resilience to infectious and non-infectious stress in pigs. The hair was sampled from the backs of 863 young and healthy Landrace x Yorkshire F1 barrows at ~40 days of age. The pigs were genotyped with a 50K SNP panel and imputed up to 650K. Previous analyses found the estimates of heritability for cortisol (CL), cortisone, DHEA, DHEA-S (DS), and the CL/DS ratio of 0.24±0.08, 0.10±0.07, 0, 0.22±0.08 and 0.33±0.08, respectively. Concerning the response to a standardized non-infectious stressor, cortisol was moderate-to-highly genetically correlated with the number and intensity of vocalizations and with the number and intensity of struggles during a 30 s back-test. Here, we present a genome-wide association study (GWAS) for the levels of these stress hormones in the hair of these pigs. The GWAS was performed with marker-based animal models using Bayesian methods implemented in the JWAS package. A 1 Mb window that explained more than 1% of the genetic variance was considered a QTL and was expanded 2 Mb up and downstream to constitute a 5 Mb QTL region (QR). Candidate genes in the QRs were identified with BioMart using the Sscrofa11.1 genome build. No QTL was identified for cortisone. A common QR on SSC2 explaining 45.3% of the genetic variance for cortisol and 1.4% for the CL/DS ratio was identified, harboring the NR3C1, SRA1, and HDAC3 genes. A QR on SSC3 explains 3.0% of the genetic variance for DHEA-S and 6.7% for the CL/DS ratio. This harbors the NCOA1 and FKBP1B genes. A QR on SSC5 explains 1.4% of genetic variance for the CL/DS ratio and harbors the ENSSSCG00000056452 and EMP1 genes. These results highlight the genetic control of response to non-infectious stress in pigs. Ongoing work focuses on evaluating the genetic relationships of hormone levels in hair from these healthy pigs with the levels in hair grown during exposure of these pigs to a polymicrobial disease challenge and with responses to the associated infectious stresses.

This study was funded by Genome Canada, PigGen Canada, Results Driven Agricultural Research, and USDA-NIFA grant #2021-67015-34562.

Automated CUT&RUN platform for high-throughput epigenomics for agricultural research

Kelsey E. Noll[1], Matthew R. Marunde[1], Danielle N. Maryanski[1], Katherine Novitzky[1], Catherine Smith[1], Allison Hickman[1], Andrea L. Johnstone[1], Keith E. Maier[1], Ellen Weinzapfel[1], Irina K. Popova[1], Bryan J. Venters[1], Keli L. Rodriguez[1], Carolina P. Lin[1], Ashton Gladney[1], Rachel Watson[1], Zachary B. Gillespie[1], Augustus Adeleke[1], Saarang Gopinath[1], Leslie Lewis[1], Hannah Willis[1], Lu Sun[1], Marcus A. Cheek[1], Matthew J. Meiners[1], James R. Bone[1], Martis W. Cowles[1], Zu-Wen Sun[1], Michael-Christopher Keogh[1]

1 EpiCypher Inc, Durham, NC

Epigenomic factors such as transcriptional signaling proteins and histone post-translational modifications (PTMs) play fundamental roles in gene regulation. Mapping the genomic localization of these features is a powerful approach to unlock new insights into areas of agricultural research such as plant development and phenotypic plasticity. Indeed, epigenomic mapping can provide more detailed insights into genetic regulation than what can be discerned by other approaches such as chromatic accessibility mapping (e.g., ATAC-seq) or transcriptomics (e.g., RNA-seq). Historically, ChIP-seq has been the leading approach for epigenomic mapping. However, ChIP-seq is plagued by shortcomings including high cell inputs, poor signal-to-noise, low sample throughput, and high cost due to deep sequencing requirements. CUT&RUN (Cleavage Under Targets and Release Using Nuclease) is a next-generation epigenomic mapping approach that delivers high signal-to-noise mapping data using a fraction of the required cells and sequencing depth compared to ChIP-seq. In CUT&RUN, antibodies are used to localize the activity of a nuclease to targeted genomic regions: DNA at antibody-targeted loci is released from the nuclei and collected for sequencing, while background genomic DNA remains in the cell. Here, we describe the development of a robust, quantitative, automation-compatible CUT&RUN workflow to enable scaled genomics studies for agricultural research. A key component of these workflows is the development and application of spike-in controls comprised of DNA-barcoded recombinant nucleosomes (dNucs), which enable rigorous in-assay technical monitoring and quality control, as well as quantitative normalization. Recently, we have successfully applied a CUT&RUN workflow with dNuc spike-in controls to profile the epigenetic landscape of nuclei from Arapidopsis thaliana, demonstrating the transformative potential of this technology to advance agricultural research. Our scalable CUT&RUN approach commoditizes the use of epigenomics for agricultural and related fields of research, enabling insights into epigenetic regulatory mechanisms underlying key processes such as plant development and environmental responses.

Bayesian attention network hybrid genomic prediction for key sugarcane traits

Chensong Chen[1], Owen Powell[2], Eric Dinglasan[1], Elizabeth Ross[1], Xianming Wei[3], Felicity Atkin[3], Emily Deomano[3], Ben Hayes[1]

1 Centre for Animal Science, Queensland Alliance for Agriculture and Food Innovation, University of Queensland, St. Lucia, Australia
2 Centre for Crop Science, Queensland Alliance for Agriculture and Food Innovation, University of Queensland, St. Lucia, Australia
3 Sugar Research Australia, Indooroopilly, Queensland, Australia

The sugarcane genome is exceptionally complex due to its high polyploidy and multi-species ancestry. Substantial dominance and epistatic effects have been demonstrated for traits such as yield and disease resistance, making genomic prediction quite challenging. Attention networks, are a class of a machine learning methods that are increasingly widely used because of their capacity to capture linear and hidden non-linear relationships in large data sets. Attention networks might be useful for genomic prediction in sugarcane, a crop with high polyploidy, multi-species ancestry and with substantial dominance and epistatic non-additive effects and gene-by-gene effects for traits such as yield and disease resistance, making genomic prediction quite challenging.

We evaluated attention networks, GBLUP methods, and Bayesian approaches (BayesR and C), as well as simple hybrids of these approaches, for accuracy of genomic prediction in a sugarcane dataset comprised 4,702 clones genotyped for 26k genome-wide single-dose markers and scored for two disease traits: smut and Pachymetra (root rot), scored on a scale of 1-9, ranging from minor to severe. For the hybrid approaches, a two-step procedure was used, where either BayesR or a GWAS was first used to select a subset of 1000 SNP (explaining the greatest proportion of additive variance) that were then fed into the modified attention network. The hypothesis here was that interaction effects would be more accurately estimated with a smaller subset of markers, leading to higher accuracy of prediction.

Of the single stage approaches, BayesR and Attention networks gave the highest accuracy of clonal prediction, when averaged across traits and cross-validations. The hybrid models gave slightly higher accuracies of clonal prediction than any other method for both traits, and also had the lowest mean square error across predictions.

In conclusion, our findings suggest that the attention mechanism holds significant potential for improving the performance of genomic prediction by enhancing its ability to capture marker interactions, particularly when combined with Bayesian methods in hybrid approaches.

Comparison of changes in SNP effects and GEBV in different training populations in simulated data

Guarav Dutta

Genome-wide association studies using SNPs aim to capture QTL effects across the genome. However, the results of GWAS vary across studies, even when traits and populations are similar. The current study aims to compare changes in SNP effects and GEBVs using different samples of a population for training. Data from 20000 individuals was simulated using the AlphaSimR package. Data for a single trait with a heritability of 0.25 and 10 chromosomes were simulated, and linkage disequilibrium was created by random crossing across 1000 generations. A total of 49,967 SNP markers were evenly placed throughout the genome. Two traits were simulated with genetic architecture controlled by 10 and 100 loci per chromosome, accounting for 100% of the genetic variation. The QTL effects were sampled from normal or from a gamma distribution. The training population was then divided into three types; all or two random halves of the individuals. Phenotypes for individuals in the last generation were masked for validation purposes. Genomic estimated breeding values (GEBV) were calculated using the single-step GBLUP method. The correlation between GEBVs of complete and reduced data was 0.84, and between each half was 0.40, while SNP effects correlations were respectively 0.75 and 0.18. Prediction accuracies varied from 0.47 to 0.38 from the complete and sampled halves, respectively. Visual inspection of Manhattan plots showed that few peak regions that explained higher variance were stable across subsamples, and experienced fewer fluctuations, while other regions often varied. The largest differences were between the two reduced random halves where only the large QTL effects were captured by the SNPs but with many false positive peaks that differ across the two subsets. Finally, the full data resulted in smaller peaks than the reduced data. The fluctuations in SNP effects in GWAS highlight the complexity of genetic influences on traits, and how sensitive those studies are to small changes in data. Understanding the factors that contribute to GWAS fluctuations may help improve the design and interpretation of such studies.

De novo genome assembly and population-based genome sequencing to assist genetic conservation of the Friesian horse breed

Marije Steensma[1], M.F.L. Derks[1], H.P. Doekes[1], M.A.M. Groenen[1], B.J. Ducro[1]

1 Wageningen University & Research, Animal Breeding and Genomics, Wageningen, the Netherlands

Corresponding author email: marije.steensma@wur.nl

In horses, genetic diversity is predominantly observed between breeds, with little variation within breeds. The Friesian horse population is distinct from other horse breeds and suffers from consequences of inbreeding. Increasingly, genomic technologies are used to aid in the conservation of such breeds. The studbooks of the two largest horse populations in the Netherlands, the Dutch Warmblood horse and Friesian horse population, are routinely collecting genotypes (70K SNP array). However, when assessing their DNA-profile, breed-specific variation will be missed because all polymorphism detections are biased towards the sequence structure of the current refence genome derived from a Thoroughbred horse. Here, we performed nanopore sequencing (R10.4, Q20+) of an F1 cross between a Dutch Warmblood horse and a Friesian horse to create two breed-specific reference assemblies using trio binning. Triocanu was used to bin the nanopore reads (80X) of the F1 cross using the short reads (50X) of the parents. Then, we used Flye and RagTag software to create and scaffold both assemblies. Scaffolded assemblies were validated with breed-specific linkage maps based on the 70K SNP array. This resulted in two high-quality, haplotype-resolved reference assemblies (contig N50 was 35Mb and 37Mb and single copy gene completeness was 99.2% and 99.3%, for the Dutch Warmblood horse and Friesian horse, respectively). Ongoing analyses, are aimed to identify structural variants between the Dutch Warmblood horse and the Friesian horse. These high quality reference assemblies form a baseline for further genetic analysis in both breeds. Further, short read sequence data (25X) of 50 Friesian horse breeding stallions and 70K SNP data of 4,000 Friesian horses will be used to assess genetic diversity and identify causal variants underlying genetic disorders in the Friesian horse breed. To genetically conserve the Friesian horse breed, a roadmap will be developed for genetic management providing an optimal balance of selection against multiple genetic disorders and conservation of genetic diversity.

Deciphering chromatin accessibility and gene expression interplay in bovine tissues

Sarem F. Khilji[1], Shangqian Xie[1], Gabrielle M. Becker[1], Katie A. Shira[1], Morgan R. Stegemiller[1], Jamin A. Smitchger[1], Janet E. Williams[1], Lauren E. Christensen[1], Ryan M. Pace[1], Denise E. Konetchy[1], Darren E. Hagen[2], Stephanie D. McKay[3], Brenda M. Murdoch[1]

1 University of Idaho, Department of Animal, Veterinary, and Food Sciences, Moscow, ID
2 Oklahoma State University, Department of Animal and Food Sciences, Stillwater, OK
3 University of Missouri, Division of Animal Sciences, Columbia, MO

The Pan-Epigenome initiative examines the epigenetic variations within and across two key species: cattle and sheep. By integrating extensive epigenetic information, a fuller understanding of the regulatory environment can be achieved, enabling researchers to delve deeper into the biological processes influencing the phenotypic variation in these animals. For the initial profiling of epigenetic diversity among cattle breeds, an intra-species comparison was performed, involving RNA-seq and ATAC-seq analysis of three tissue types (liver, skeletal muscle, and prefrontal cortex) from four 3-month-old Angus steers. Differential gene expression analysis (|log2(fold change) | > 2.5 and p <0.01) identified 4,839 genes in brain vs liver, 4,151 in brain vs muscle, and 3,664 in liver vs muscle, highlighting the unique functional characteristics of these tissues. In addition, differential transcript usage (DTU) analysis detected over 400 genes in each comparison that demonstrated notable differences in the proportion of expressed isoform composition, with a marked number of genes identified as DTU's uniquely in each tissue type. Furthermore, 1,876 genes in the brain cortex, 925 in muscle, and 1,458 in the liver were identified as tissue enhanced. Analyses using ATAC-seq data demonstrated that a significant portion of the enhanced genes, specifically 79% in brain, 83% in muscle, and 82% in liver, also exhibited open chromatin peaks, supporting a link between RNA expression and chromatin accessibility. Analysis of chromatin accessibility peak distribution across genomic features highlighted a predominance of peaks within intergenic, intronic, and promoter regions in all three tissue types, aligning with the known biological significance of these regions, which harbor elements essential for the regulation of gene expression. A study of peaks near transcription factor-binding sites revealed a high density near transcription start sites, particularly within 1 kilobase (kb), with a gradual decrease as distance increased. Despite this, binding was detected up to distances greater than 100 kb, suggesting the involvement of long-range regulatory interactions in gene control. This research not only aims to illuminate the intricate relationships between genetic, epigenetic, and transcriptomic factors in these species, but also to advance our understanding of how these elements contribute to the tapestry of phenotypic diversity.

Developing an app for statistical analysis of groups of experiments, with applications in plant breeding

Brenda Monis Moreno[1], Pedro Henrique Massaro de Silveira[1], Vitória Bizão Murakami[1], Cristiane Hayumi Taniguti[2], Jenifer Camila Godoy dos Santos[1], Liana Jank[3], Mateus Figueirêdo Santos[3], Antonio Augusto Franco Garcia[1]

1 Department of Genetics, Luiz de Queiroz College of Agriculture - University of São Paulo, Piracicaba, São Paulo, Brazil
2 Texas A&M University, College of Agriculture and Life Sciences, College Station, TX
3 Brazilian Agricultural Research Corporation (Embrapa) Beef Cattle, Campo Grande, Mato Grosso do Sul, Brazil

In plant breeding, the predominant use of the "randomized complete block" design in genetic evaluations is noteworthy, especially in tropical forage crops, where superior genotypes are sought through local control. Brazil, a leading beef exporter with the most significant commercial beef herd, critically depends on pasture for its national production. Despite its success, the commercial forage supply is limited, with Brachiaria brizantha cv. Marandu and Panicum maximum cv. Mombaça dominating. The analysis of Panicum maximum shows a great genetic potential that varies according to the environment and climatic conditions, such as rain and drought. Therefore, evaluating different varieties under different conditions over the years is essential for crop improvement, providing insights into genotype-by-environment interactions (GEI). In this context, the availability of efficient computing resources is essential for researchers. In this context, the goal of this study was to develop a computer app that would allow breeders to analyze their data more quickly and efficiently, facilitating the conclusions regarding the impact of GEI on the traits assessed. As an example, the app tested leaf dry matter (LDM) phenotypic data collected from 23 genotypes of P. maximum. This study, carried out by Brazilian Agricultural Research Corporation (Embrapa) Beef Cattle, used a randomized complete block design in five locations, each with up to seventeen harvests. To deal with incomplete data and to ensure greater flexibility, the analysis was performed using the mixed model according to the experimental design implemented. Genetic and residual correlations between environments could be modeled by fitting different variance-covariance structures for random effects. This model is applied separately for each harvest. The analyses are performed through the "DesignGen" application, allowing the researcher to interpret the GEI. Therefore, the development and implementation of the DesignGen app represents a significant advancement in the GEI analysis and the improvement of forage breeding programs. The public release of this tool provides breeders with a more efficient and accurate way to analyze data and facilitate the identification of superior genotypes.

Development of functions in package onemap/r to perform grouping of markers into linkage groups during linkage maps construction

Marina Cunha Bernardes[1], Cristiane Hayumi Taniguti[2], Heitor Calux Mendonça[1], Antonio Augusto Franco Garcia[1]

1 University of Sao Paulo, Luiz de Queiroz College of Agriculture, Department of Genetics, Piracicaba, Sao Paulo, Brazil
2 Texas A&M University, College of Agriculture and Life Sciences, College Station, TX

Genetics linkage maps are meaningful tools for genetics studies, enabling multiple applications in breeding programs, such as quantitative trait loci (QTLs) analyses, synteny studies, and genome assembly. OneMap/R software, a package within R programming environment, facilitates the construction of genetic maps for several types of crosses and molecular markers. However, continuous improvements are implemented in its code to maintain the significance and relevance of the software. In order to enhance the efficiency of clustering molecular markers into linkage groups in OneMap/R, the objective of this research is to test k-means and k-Nearest Neighbors (KNN), statistical clustering algorithms, in the development of new functions for the software, respectively, group_kmeans() and group_knn(). For the development of these functions, RStudio was the tool used for programming processes, dataset simulation, data analysis, and evaluation of results obtained from algorithm tests. The tests aimed to explore clustering efficiency and the initialization effect of using "guides" for clustering, such as genomic information, published maps, and so on. As a result, function group_kmeans() performs efficiently, reaching an accuracy of 100%. Although, it is dependent on initial positions guides, once the probability of an erroneous final clustering increases when the initial clustering is more divergent from the correct positions, since in absence of data guides, clustering accuracy was only 50%. On the other hand, function group_knn() showed high efficiency in clustering the markers, even with limited guide data, once it achieved an accuracy of 100% for all tested ks and percentages of markers, which suggests that even with few initial genetic information and a low number of recombination fractions selected the function is highly effective. In conclusion, the kNN and k-means algorithms have proven to be efficient clustering tools. Currently, a new function group_pca(), based on principal component analysis, is in the final stages of development.

Discovery of genetic diversity of regulatory elements responsive to viral infections

Chaelynne Lohr[1], Alix Booms[2], Andrew Borgman[3], Casey Droscha[1,4]

1 CentralStar Cooperative, Lansing, MI
2 Van Andel Institute, Grand Rapids, MI
3 Borgman Consulting Group LLC, Alma, MI
4 Department of Animal Science, College of Large Animal Clinical Sciences, Michigan State University, East Lansing, MI

Advancements in genetic selection revolutionized breeding decisions in the bovine industry, enabling producers to prioritize economically significant traits like milk production. Despite progress, addressing chronic diseases like Bovine Leukemia Virus (BLV) remains challenging, with a 50% national prevalence. Bovine genotyping methods rely on SNP chip data and imputation to predict haplotype states at selected loci. However, these techniques cannot accurately genotype the Bovine Leukocyte Antigen (BoLA), the bovine Major Histocompatibility Complex (MHC), due to its highly complex and polymorphic nature. Furthermore, research indicates that many disease-associated genetic variations are within non-coding regions, especially regulatory elements, not discerned by current genotyping techniques. Despite the popularity of SNP-based genomic selection among producers, it's vital to consider its impact on bovine genetic diversity in the US. A fundamental principle of population biology emphasizes the importance of genetic heterogeneity for robust population-level immunity, urging a comprehensive approach to understanding genetic diversity within the bovine genome. Yet, there exists untapped genetic data within the bovine genome that requires specialized genotyping and the study of immune-driven phenotypes to understand their role in immunity. This data could offer vital insights into genetic resilience to diseases, an unexplored area of the dairy industry. Recently, our research noted distinct BLV disease phenotypes, where certain animals maintain high levels of BLV proviral DNA (BLV susceptible), while others, despite being seropositive, exhibit negligible detectable virus levels for extended periods (BLV resilient). Investigating BoLA genes, directly involved in pathogen recognition and response, from each phenotypic group revealed that specific alleles are associated with BLV resilience and highlighted the limited BoLA allele diversity within the Midwest dairy population. To expand our findings, we aim to determine whether progeny from bulls carrying BLV-resilient alleles inherit the BLV-resilient phenotype and to delve deeper into understanding the phenotypic consequences of reduced diversity within immune-related transcriptional regulatory regions. Our long-term goal is the discovery of novel genetic elements implicated in bovine immune response, with the overall objective of evaluating how these regions have changed due to intense genetic selection for milk production and understand the impact of these changes on the offspring of contemporary elite sires.

Efficient Estimation of Genetic Parameters for a Resilience Indicator as Environmental Variance

Noelia Ibáñez-Escriche[1], Cristina Casto-Rebollo[1], Alejandra Alvarez Munera[2], Daniela Lourenco[2], Ignacy Misztal[2], Matias Bermann[2]

1 Department of Animal Science, Universitat Politècnica de Valencia, Valencia, Spain
2 Department of Animal and Dairy Science, University of Georgia, Athens, GA

Understanding how animals adapt to environmental factors is crucial for promoting a more sustainable livestock system. Phenotypic indicators rooted in repeated records, such as root mean square error (RMSE), logarithmic deviations from expected patterns (LnVar), or residual variance (VE), have been shown to be good indicators of resilience. However, the estimation of the genetic parameters of these indicators usually requires prior correction for confounding factors, which assumes that the residuals are true data. This can lead to bias in the estimation of genetic parameters, especially in the genetic correlations with production traits. Bayesian and frequentist methods based on the double hierarchical general linear model (dhglm) have been developed to model VE. These approaches account for environmental and genetic control in mean and variance, ensuring an accurate estimation of genetic parameters and positioning it as a promising trait for assessing animal resilience. However, Bayesian models have a high computational demand, the frequentist software packages are not user-friendly, and the convergence is sometimes unstable. A modified iterative reweighted least square approximation to solve dhglm models has been implemented in the blupf90 family software. This study aimed to accurately calculate the genetic parameters of VE, using the dhglm tool of the blupf90 family. For this purpose, we used simulated data and data from daily feed intake (DFI) records of pigs used in previous studies to calculate RMSE, LNVar, and VE resilience indicators. Our results showed that the new method allows the estimation of covariance variances and all effects in at least 30 times less time than Bayesian analysis. Furthermore, we propose a genome-wide association study (GWAS) incorporating this model to identify the genomic regions specifically associated with VE of DFI and distinguish it from that associated with trait mean.

Evolution and function of receptor kinases in arbuscular mycorrhizal symbiosis

Chai Hao Chiu[1], Leo Minglu Li[1], Serena Qiao[1], Jonathan Tsoi[1], Pablo Hernandez-Martinez[1], Ruth Bates[2], Tom Lawrenson[3], Wendy Harwood[3], Emma Wallington[2], Giles Oldroyd[1], Uta Paszkowski[1]

1 University of Cambridge, Crop Science Centre, Cambridge, United Kingdom
2 University of Cambridge, National Institute of Agricultural Botany (NIAB), Cambridge, United Kingdom
3 John Innes Centre, Norwich Research Park, Norwich, United Kingdom

Evolution of symbiosis with arbuscular mycorrhizal fungi (AMF) was one of the key innovations that enabled the successful conquest of land plants. In this mutually beneficial symbiosis, fungi-delivered mineral nutrients are exchanged for plant photosynthates in elaborate, highly-branched arbuscules. This symbiotic association is important in nature, occurs in natural and agricultural ecosystems, and is a crop trait targeted for engineering and improvement. When engaging with its microbiome, plants detect micro-organisms and non-self molecules via a complex immune surveillance system including receptor kinases located on the plasma membrane. However, we do not fully understand how the early detection of the symbiotic fungus via receptor kinases lead to appropriate signal transduction, as well as the types of signals and receptors involved in the processes. At the same time, signalling activated by the fungus also reprogrammes plant root development to produce more lateral roots, resulting in increased surface area for symbiotic interactions.
We discovered that a set of plant receptors are involved in symbiosis establishment, as well as a conserved signal-receptor system in activating lateral organs in response to chitin molecules. Tracing the evolution of these receptors in the green lineage also revealed insights on the origins and deep conservation of their symbiosis signalling functions early in land plant evolution. I will also discuss ongoing work to elucidate the signal transduction cascades downstream of the receptor kinases.

Genetic architecture of root-knot nematode resistance in common beans

Talissa Oliveira Floriani Z. Souza[1,2], Henrique Castro Gama[2], Guilherme Alexandre Luz da Costa[2], Bruna Marques Moreno[2], Willian Giordani[3], Alisson Fernando Chiorato[4], Luis Eduardo Aranha Camargo[5], Maria Lucia Carneiro Vieira[2], Antonio Augusto Franco Garcia[2]

1 University of Illinois Urbana-Champaign, Department of Crop Sciences, Urbana, IL
2 University of São Paulo, Department of Genetics, Piracicaba, São Paulo, Brazil
3 Corteva Agriscience, Toledo, PR, Brazil
4 Campinas Institute of Agronomics, Campinas, São Paulo, Brazil
5 University of São Paulo, Department of Phytopathology and Acarology, Piracicaba, São Paulo, Brazil

The common bean (Phaseolus vulgaris L.) is a crop used for direct human consumption. The grain covers essential micronutrients and valuable protein content, thus contributing to food security in developing countries. However, attacks by root-knot nematode (RKN - Meloidogyne incognita) threaten common bean cultivation, causing substantial yield losses. Undoubtedly, crop resistance is an outstanding approach for suppressing nematode infection. Contrasting genotypes were identified and crossed to unveil the genetic architecture of typical bean response to RKN (race 3 of Meloidogyne incognita). The segregating population (F2), which consisted of 388 individuals, was genotyped using GBS (genotyping-by-sequencing), and a customized high-throughput phenotyping approach was developed to acquire trait data for a subset of 200 F2:3 families. The traits egg mass (EM), root-galling index (RG), and root dry mass (RM) were evaluated over time within greenhouse conditions under a completely randomized design with ten replicates. A total Linkage map resulting in 954 SNPs assigned to 11 linkage groups totaling 1,687 cM was used as a basis for Composite Interval Mapping (CIM) and Multiple Interval Mapping (MIM), identifying four major QTLs (Pv03, Pv05, Pv08 and Pv10). The selected model was used to calculate the genotypic values of individuals, with the marker-assisted selection (MAS) approach, arriving at a list of the top 10 genotypes for use in MAS. The correlation between observed and predicted values was 0.72, which is considered high; the result shows the relevance of the model. Candidate genes were identified in response to GI, with domains related to WRKY and MAPK (Mitogen-Activated Protein Kinase) cascades. This work represents a significant step in understanding the genetic architecture of RNK resistance in the common bean. It sets the stage for implementing MAS in a breeding population against pathogen resistance in the bean family.

Genetic insights into selection for heat tolerance in feral cattle in Northern Australia

Jigme Dorji[1], Antonio Reverter[1], James Kijas[1], Laercio Porto-Neto[1]

1 Commonwealth Scientific and Industrial Research Organisation (CSIRO) Agriculture & Food, Livestock & Aquaculture, Brisbane, Queensland, Australia

The feral cattle population in Northern Australia originated from escaped or abandoned domesticated cattle, leading to animals adapted to the challenging tropical climatic conditions of the region (e.g. extreme heat and humidity at times). The adaptation to extremely hot tropical conditions through natural selection is one of the interesting attributes of feral cattle and offers a unique opportunity for genetic and genomic studies to understand the underlying biology of heat tolerance. While these feral cattle are believed to be predominantly crossbreds with Brahman type, it is known that in the early days of the agriculture expansion in the North, Shorthorn cattle were introduced in the region and some of their progenies are believed to have survived and adapted. In this study, we profiled DNA of feral cattle using a 50,000 SNP markers array from three populations in northern Australia. The genetic analysis revealed a significant genetic diversity with an overall mean nucleotide diversity (π) of 0.39. The feral cattle were composed of varying blood levels of indicine (0-96%) derived mainly from the Brahman breed, and taurine (0-100%) consisting of Shorthorn, Senepol, and Tropical Composite types. There was also significant genetic differentiation between the feral cattle populations (FST 0.18). In the next phase of our study, we intend to conduct genetic comparison between northern feral cattle with a higher proportion of taurine ancestry and animals with similar ancestry from the southern regions to uncover signatures of selection and identify genomic regions and genes that play a role in tropical adaptation. Altogether, the findings of this study will enhance the understanding of the genetic factors influencing heat tolerance and climate adaptation in cattle. This knowledge is expected to drive the development of breeding strategies aimed at improving heat resilience and climate adaptability in livestock populations.

Genetic relationships among immunoassay traits

Vishesh Bhatia[1], Frederic Fortin[2], Michael K. Dyck[3], Catherine J. Field[3], Claire Rogel-Gaillard[4], Fany Blanc[4], Graham S. Plastow[3], Jack C.M. Dekkers[1]

1 Department of Animal Science, Iowa State University, Ames, IA
2 Centre de Développement du Porc du Québec, Inc., Quebec City, Quebec, Canada
3 Livestock Gentec, Department of Agricultural, Food, and Nutritional Science, University of Alberta, Edmonton, Alberta, Canada
4 Université Paris-Saclay, INRAE, AgroParisTech, GABI, France

Infectious diseases take a severe toll on the swine industry, resulting in large financial and production losses each year. Breeding for disease resilience, defined as an animal's ability to maintain production performance under disease, has gained traction recently as an approach to help reduce the impact of infectious disease. To further investigate this, a wean-to-finish polymicrobial natural disease challenge model was established, in which, using a continuous flow system, batches of 60 or 75 young healthy Landrace × Yorkshire barrows from one of seven breeding companies of the PigGen Canada research consortium were entered into a quarantine nursery prior to their entry in a nursery/finisher that was seeded with multiple pathogens. Prior to their entry into the challenge, blood samples collected at ~26 days of age were evaluated for immune cell proportions and their phagocytic activity (n=985), and for cytokine concentrations and mean fluorescence intensity (n=1524) following stimulation with Lipopolysaccharide (LPS) or Phorbol myristate acetate and Ionomycin (PI), as well as prior to stimulation. The objective of this study was to investigate the underlying biological and genetic relationships among these immunoassay traits. Genetic parameters, including heritabilities and genetic correlations were estimated using GBLUP in ASReml 4.2, with the fixed effect of batch and assay date, the covariate of entry age, and random effects of pen by batch, litter, and additive genetics, using a genomic relationship matrix derived from genotypes of 650K single nucleotide polymorphisms. Results showed moderate heritability estimates for immune cell proportions (average = 0.36; range = 0.22-0.42), the % of phagocytizing cells for each cell type (0.32; 0.19-0.48), and blood stimulation measures (0.08; 0-0.27), but low estimates for mean fluorescence intensity of the phagocytizing cells of each type (0.2; 0.05-0.3). Downstream analyses will involve incorporation of principle component and factor analyses to explain the observed genetic relationships between the immunoassays. This project was supported by Genome Canada, Genome Alberta, PigGen Canada, and USDA-NIFA grant #2017-67007-26144.

Genomic imprinting in livestock species: Challenges of long-read sequencing

Julie Demars[1], Jean-Noël Hubert[1], Nathalie Iannuccelli[1], Sophie Leroux[1], Céline Vandecasteele[3], Paul Terzian[2], Thomas Faraut[1], Christophe Klopp[2], Cécile Donnadieu[3]

1 GenPhySE, INRAE, Université de Toulouse, INPT, ENVT, Castanet-Tolosan Cedex, France
2 GenoToul Bioinformatics Facility, Université Fédérale de Toulouse, INRAE, Castanet-Tolosan Cedex, France
3 GeT-PlaGe, Université Fédérale de Toulouse, INRAE, Castanet-Tolosan Cedex, France

Genomic imprinting is an epigenetically-regulated process of central importance in mammalian development and evolution. It involves multiple levels of regulation, with spatio-temporal heterogeneity, leading to the context-dependent and parent-of-origin specific expression of a small fraction of the genome. Genomic imprinting studies have therefore been essential to increase basic knowledge in functional genomics, evolution biology and developmental biology, as well as with regard to potential clinical and agrigenomic perspectives. However, suitable molecular and computational tools might be challenging especially in livestock species in which knowledge on genomic imprinting is still sparse.
Here, we propose an insight of two technologies (targeting sequencing through short reads vs. genome-wide sequencing through long reads) to detect parent-of-origin methylation with a focus on germline specific differentially methylated regions (DMR). Both studies led in pigs using reciprocal crosses showed a higher rate of genotyping errors and a higher variability of CpG methylation rate with the long-read sequencing technology than the capture methyl-seq technology developed by our group. However, when both technologies are concordant, information is more exhaustive using the long read technology thanks to the unambiguous haplotype identification of the two parental origins. Although, analyses need to be deepened, we identified orthologous paternal and maternal DMR as well as pig specific parental DMR. Our results highlight the necessity of complementary sequencing technologies to bring new insights for a better understanding of genomic imprinting mechanisms in livestock species. Moreover, the validation of a dedicated molecular tool in pigs focused on imprinted regions paves the way to novel studies to evaluate the contribution of genomic imprinting mechanisms in the genetic architecture of major agronomic traits.

Genomic information on X chromosome: Assessing the imputation accuracy

Renata Veroneze[1,2], Ching-Yi Chen[3], Justin Holl[3], Daniela Lourenco[2]

1 Universidade Federal de Viçosa, Animal Science, Viçosa, Minas Gerais, Brazil
2 University of Georgia, Department of Animal and Dairy Science, Athens, GA
3 The Pig Improvement Company, Genus plc, Hendersonville, TN

The X chromosome (CHRX), often overlooked in genomic predictions, may have important contributions to genetic variance. Using CHRX information in genomic predictions may increase the breeding value accuracy. Imputation can be challenging in CHRX as it has pseudoautosomal and non-pseudoautosomal regions that may hamper imputation accuracies. Thus, knowledge of CHRX imputation quality is fundamental for genomic predictions. Data from 51,038 pigs genotyped with 41,963 SNPs, including 1494 on CHRX, were available. The quality control was done using preGSF90 software. After quality control, 50,932 animals and 41,548 SNPs, of which 1445 were on CHRX, remained for the analysis. From the complete marker set it was carried an independent random selection of 1,000, 3,000, and 10,000 markers. It will be considered as the low-density panels (LD). These panels included 31, 90, and 314 markers in CHRX, which represent 2.15%, 6.23%, and 21.73% of the total CHRX markers remained, respectively. The genotyped animals were randomly divided into five independent groups to carry out a five-fold cross-validation. Imputation accuracy (R^2) and concordance rate (CR) were computed to evaluate imputation quality. The imputation was done using FIMPUTE. Imputation accuracy and CR increased with the number of SNPs for the autosomes and CHRX. The CHRX accuracies were 0.78 and 0.93 when using 1,000 and 3,000 markers, respectively. These values are lower than the averaged accuracies computed using all autosomes in the same scenarios, 0.84 and 0.95, respectively. Similar results were observed for the concordance rate. When using 10,000 markers, no differences in the imputation were observed for autosomes and CHRX. In this scenario, R^2 of 0.98 and CR of 0.99 were found for both. We also found lower R^2 and CR at the start and the end of CHRX in 1,000 and 3,000 scenarios, compared to the mid-region of CHRX. Similar patterns for accuracy and CR in relation to minor allele frequency (MAF) for autosomes and CHRX were observed. Overall, imputing SNPs in CHRX is practical. Different LD panels should be tested to establish the minimum number of markers to ensure accurate imputation.

Hierarchical genomic prediction models as a framework for biological discovery

Owen Powell[1,2], Greg McLean[1,2], Jason Brider[1,2], Frank Technow[3], Tom Tang[3], Radu Totir[3], Carlos D Messina[2,4], Graeme Hammer[1,2], Mark Cooper[1,2]

1 The University of Queensland, Queensland Alliance for Agriculture and Food Innovation (QAAFI), Centre for Crop Science, Brisbane, Queensland, Australia
2 The University of Queensland, ARC Centre of Excellence for Plant Success in Nature and Agriculture, Brisbane, Queensland, Australia
3 Corteva Agriscience, Johnston, IA
4 University of Florida, Department of Horticultural Sciences, Gainesville, FL

Methods for prediction of crop yield and resource use that account for the effects of genotype (G), (M) management, (E) environment, and their interactions (GxExM) create many opportunities for enhancing the productivity and sustainability of agricultural systems. Crop growth models, symbolic networks of expert knowledge on trait-environment relationships, have a history of application for crop management decision support. Genomic prediction models, sub-symbolic networks of genome-phenotype relationships, have a history of application in breeding to enable the prediction of crop traits for genotypes. Both classes of prediction tools are used to optimise the design of crop improvement programs. However, domain-specific optimisation strategies can limit the exploration of possible (GxM) solutions when compared to optimisation strategies with the potential to explore the full GxExM state space required to unlock agricultural productivity and sustainability improvements. The development of hierarchical genomic prediction models, integrating symbolic and sub-symbolic networks, provides a single end-end framework to predict a broader range of yield and resource use properties of complex G×E×M interactions across multiple levels of the genome-to-phenome (G2P) biological hierarchy. We present a novel hierarchical G2P prediction methodology, linking the APSIM sorghum crop growth model with the BayesA genomic prediction algorithm (APSIM-G2P). APSIM-G2P predictive accuracies for grain yield are compared to a standard BayesA alone G2P prediction approach in a simulated sorghum multi-environmental trial dataset for the Australian Target Population of Environments. Finally, we explore and discuss nascent opportunities for hierarchical G2P models to deconvolute genetic interaction effects and improve the identification of biological discoveries at the genomic and component trait levels of the biological hierarchy.

Identification of genomic region associated with the causal quantitative trait locus (QTL) of the spontaneous haploid genome doubling (SHDG) trait in Ames panel by genome-wide association study (GWAS)

Mercy Fakude[1], Ursula K. Frei[1], Thomas Lubberstedt[1]

1 Iowa State University, Department of Agronomy, Ames, IA

The doubled haploid (DH) approach is a promising alternative to traditional self-pollination for inbred line development, primarily because it decreases the time taken to obtain homozygous lines. Currently, DH technology is widely dependent on chemical treatment (colchicine) to induce haploid genome doubling and, subsequently, male fertility. These chemicals can be harmful to humans and plants and have typically recorded a doubling rate of about 10 to 30% due to the recalcitrance of some temperate genetic materials. Thus, genome doubling remains a bottleneck to large-scale DH production. Although to a limited extent, spontaneous haploid genome doubling (SHGD) and male fertility of maize haploids (without chemical treatment) have been studied and recorded doubling rates ranging from less than 5% to greater than 50%. Recently, there has been increased interest to omit chemical treatment and rely on SHGD as it would eliminate the use of colchicine for genome doubling and make it possible to directly plant a haploid seed in the field and avoid greenhouse costs. A significant breakthrough was the discovery of a maize line A427 with SHGD and a major QTL mapped to chromosome 5. The introgression of this major QTL to elite germplasm may overcome the need to use colchicine for inbred line development. Therefore, it is necessary to identify additional sources of major QTL for SHGD in maize as they may (1) simplify managing SHGD in different heterotic groups and (2) allow simplified introgression procedures into elite germplasm by stacking respective SHGD QTL. This genome-wide association study (GWAS) aims to identify additional sources of major QTL controlling SHGD.

Impact of genomic selection for growth on feet and leg structure in Angus cattle

Mitzilin Zuleica Trujano Chavez[1], Jorge Hidalgo[1], Daniela Lourenco[1], Kelli Retallick[2], Andre Garcia[2], Ignacy Misztal[1]

1 University of Georgia, Department of Animal and Dairy Science, Athens, GA
2 American Angus Association, Angus Genetics Inc., Saint Joseph, MO

We aimed to examine genetic parameter changes in feet and leg structure, and growth traits over time in Angus cattle. Additionally, new predictivity-based formulas were tested for estimating heritability and genetic correlations. The analysis was conducted using the blupf90+ family of programs. To estimate variance components (VCE), sampled data (2011-2022) from the American Angus population included 311K growth traits (GT) records [birth (BW), weaning weight (WW), and post-weaning gain (PWG)], ~70K feet traits (FT) records [claw set (CS) and feet angle (FA)], and 142K genotypes. The dataset was split into five-year intervals. Since FT ranged from 5-9, and 5 is ideal, a negative correlation is desired between FT and GT. The correlation FA-PWG, changed in an undesirable way over time, as it was -0.15 in the first interval, and it was 0.02 in the last interval. The correlation FA-WW is becoming more favourable over time, shifting from 0.05 in the first interval to -0.18 in the last interval. The GT selection process shifted the correlations between CS and GT in the desired direction, changing from positive (~0.1) to almost zero in the more recent two intervals. To test the predictivity-based formulas, Genomic Estimated Breeding Values (GEBV) were estimated using GT data from 6.3M animals born from 2000 to 2022. The data was divided into four slices, each with a 2-year validation period. Over time, heritabilities remained constant: ~0.3 for BW and WW and ~0.2 for PWG. Correlations between GT also remained constant over time, approximately at 0.45 for BW-WW, 0.3 for BW-PWG, and 0.75 for WW-PWG. Although the formulas show promising results for heritabilities, interpreting the outcomes of genetic correlation estimation remains unclear, needing further tests. While certain correlations between FT and GT suggest a positive impact of genomic selection over time, there are also some undesirable outcomes. This emphasizes the need for a multi-trait genomic selection approach to mitigate genetic antagonism. The adoption of predictivity-based formulas not only reduces the time required for genetic parameter estimation but also enables the utilization of complete data information instead of sampled data, as is currently done

Indirect predictions for indicine cattle breeds based on a multi-breed genomic evaluation

Marisol Londoño-Gil[1,2], Jorge Hidalgo[2], Fernando Baldi[1], Daniela Lourenco[2]

1 Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal, SP, Brazil

2 University of Georgia, Department of Animal and Dairy Science, Athens, GA

Indirect predictions (IP) are used for young genotyped animals that lack phenotypes or are from commercial herds. These animals are often excluded from official evaluations to avoid decreasing accuracy and increasing inflation and bias in genomic breeding values (GEBV). In Brazil, information is scarce for breeds like Brahman, Guzerat, and Tabapua. Indirect predictions based on a larger reference population could enhance genomic selection accuracy on these breeds. This study aimed to compute IP for young genotyped Nellore, Brahman, Guzerat, and Tabapua animals in the presence of metafounders (MF) to model genetic differences across breeds. Metafounders are pseudo-individuals acting as proxies for animals in the base population. Records from the four breeding programs of the National Association of Breeders and Researchers (ANCP—Ribeirão Preto, SP, Brazil) were used. Data included pedigree (4.2M), phenotypes (329K), and genotypes (63.5K) across all breeds. The traits analyzed were adjusted weight at 210 (W210) and 450 (W450) days of age and the scrotal circumference at 365 days of age (SC365). Single-breed and multi-breed analyses were performed. Indirect predictions were derived as the sum of the SNP effects weighted by the gene content using three reference populations: multi-breed with MF, Nellore, or within-breed. The robustness of IP was evaluated as the correlation between IP and GEBV of young, genotyped animals when their phenotypic information was omitted from the official evaluations. The predictive ability was measured as the correlation between phenotypes adjusted for fixed effects (Y*) and IP or GEBV. The results indicate that IP works with the MF approach, as indicated by its high correlation (0.99) with GEBV and similar predictivity. Moreover, combining breeds increased the accuracy of IP, mainly when the number of genotyped animals in a single breed was small. These findings suggest that when young genotyped animals are not included in an official multi-breed evaluation, robust IP can be obtained whether metafounders are used or not, independent of the breed. This helps obtain fast genomic predictions for young animals without overwhelming the evaluation system.

Inhibition of nitrification in maize, an innate function of plants as an alternative in optimizing the use of nitrogen fertilizers in soil

César Petroli[1], Guntur Subbarao[2] Juan Burgueño[1], Tadashi Yoshihashi[2], Hui-hui Li[1,3], Jorge Franco[4], Kevin Pixley[1]

1 International Maize and Wheat Improvement Center (CIMMYT), Texcoco, Estado de México, México
2 Japan International Research Center for Agricultural Science, Tsukuba, Ibaraki, Japan
3 Institute of Crop Sciences, Chinese Academy of Agricultural Sciences (CAAS), Beijing, China
4 Departamento de Biometría y Estadística, Facultad de agronomía, Universidad de la República, Paysandú, Paysandú, Uruguay

In contemporary agricultural practices, the expanded application of nitrogen fertilizers on cropland has given rise to multifaceted challenges, particularly greenhouse gas emissions and water pollution. Biological Nitrification Inhibition (BNI) is the intrinsic capacity of certain plants to prevent the formation of nitrates in the rhizosphere by suppressing nitrifying bacteria in soils. It can be manifested through the release of two distinct categories of BNIs from plant roots systems, hydrophobic BNIs and hydrophilic BNIs. Although the manifestation of this biological function has been discovered in maize (Zea mays L.), the comprehensive characterization of its components has had limited progress. Recently, we developed a Genome-Wide Association Study (GWAS) analysis, using a core set of 250 CIMMYT maize lines (CML), to evaluate its hydrophobic-BNI-capacity. The results highlighted substantial variation among CMLs, with specific lines exhibiting a tenfold greater hydrophobic-BNI-capacity than those with lower capacity. Eighteen significant markers were identified through this analysis, three associated with specific BNI-activity (SBNI) (i.e. BNI-activity released g-1 root drywt. d-1), four with BNI-activity per plant (BNIPP); and another ten markers are common to both SBNI and BNIPP. In addition, a marker was related to the release of zeanone, a BNI-component that has greatest influence on the hydrophobic-BNI-capacity in maize. Our research also revealed six candidate genes correlated with nitrogen use efficiency (NUE), targeting the genetic regions responsible for some of the nitrogen-related traits. Furthermore, one of the significant SNPs has exposed the gene AMT5 on chromosome 10, recognized for its role in regulating the ammonium transportation in maize roots. We have already developed over 400 doubled haploid lines resulting from crosses between high and low BNI-activity CMLs, which will be evaluated soon for this trait. These lines are expected to facilitate the establishment of a specific set of markers, a potential tool to be used in breeding processes targeted at enhancing BNI in maize. This research has significant potential to project a reduction in nitrogen losses from maize production systems in future, which would confer important economic and environmental benefits. The application of such advances is important to drive progress in the agricultural sector, particularly in resource-limited countries.

Large-scale, multibreed genomic evaluation for fertility traits in U.S. dairy cattle using single-step genomic best linear unbiased prediction (BLUP)

Joe-Menwer Tabet[1], A. Legarra[2], I. Misztal[1], D. Lourenco[1]

1 Department of Animal and Dairy Science, University of Georgia, Athens, GA
2 Council on Dairy Cattle Breeding, Bowie, MD

The official U.S. dairy genetic evaluation involves a multistep genomic process, initially estimating pedigree breeding values (e.g., using BLUP), then SNP effects, direct genomic predictions, and genomic breeding values (GEBV) using a selection index. Single-step Genomic BLUP (ssGBLUP) streamlines the computation of GEBV by utilizing all available data in a single run. This study evaluates the official U.S. dairy genetic evaluation's performance for fertility traits employing ssGBLUP and BLUP frameworks. Working with fertility traits is challenging because of low heritability and sparse recording. The pedigree dataset comprises 93 million animals with over 2 million genotypes across all breeds in the Official U.S. multibreed genomic evaluation. Phenotypic records from 1960 to 2023 for Cow Conception Rate (CCR), Early First Calving (EFC), Heifer Conception Rate (HCR), and Daughter Conception Rate (DPR) were analyzed. The four-trait multibreed model integrated Unknown Parent Groups (UPG) and Metafounders (MF) defined by breed, year of birth, and selection pathway, resulting in 417 genetic group definitions. Evaluation methods included bias and dispersion assessments using the Linear Regression method and the Pearson correlation for (G)EBV of each analysis. Validation was performed on Jersey (JE) and Holstein (HO) breeds for CCR and DPR, and the (G)EBV of validation animals for each trait were compared. Genetic trends between ssGBLUP and BLUP with UPG or MF were examined. Bias was identified in evaluations with ssGBLUP and BLUP with UPG or MF. Notably, ssGBLUP shows better dispersion levels with MF in HO and JE compared to other models. High correlations were observed in ssGBLUP with MF for HO and ssGBLUP with UPG for JE. Genetic trends showed consistency across all models for CCR and DPR, with a decrease in trend from 1960 to 2000 and positive trends between 2000 and 2020, with ssGBLUP MF presenting the smoothest trend. The results underscore the feasibility of ssGBLUP for fertility traits in large datasets, with an emphasis needed for re-evaluating the models to minimize estimation bias.

Optimizing the multi-environment trials in the U.S. Rice Belt via Smart-Climate-Soil Prediction-based-models and economic importance

Roberto Fritsche-Neto[1], Melina Prado [1], Adam Famoso[1], Roonie Levi[1], Kurt Guidry[1]

1 Louisiana State University Agricultural Center, Rice Research Station, Rayne, LA

The phenotypic evaluation is one of the most expensive parts of a breeding program. Therefore, defining the best trade-off between the number and allocation of trials and the broad-sense heritability (accuracy) is crucial. Furthermore, maximizing the correlation between the multi-environment trials and the target population of environments increases the probability of success of newly released varieties. Thus, we aimed to optimize the LSU rice multi-environment trials (MET) via smart-climate prediction-based models using soil, historical weather, and yield data. For that, we used soil information and ten years of historical weather data during the growing seasons (2010 – 2019) for all the USA counties that significantly produce rice (77 counties/parishes representing 98% of the US rice). Then, this information was translated into rice response based on the following cardinals: Tbase1 = 12, Tbase2 = 24, Topt1 = 33, Topt2 = 37, and crop stages: 0, 45, 60, 75, 90, 105, maturity. To reduce the number of environmental covariates (EC), we first eliminate those highly correlated EC (>0.98) and then apply a supervised algorithm for feature selection. In this step, for the training and validation sets, we used two independent years of data (2021-22), composed of 43 genotypes evaluated for grain yield in 18 representative locations in the Southern USA. Based on the final set of ECs, we estimate the environmental relationship matrix (ERM) and clusters. Finally, considering the economic importance of each county and the environmental group to which they belong, we weigh the trial's allocation. Our findings show that eight ECs explained 58% of grain yield variation across sites and 53% of the "real" GxE. Moreover, it is possible to reduce 2/3 the number of trials without significant loss in accuracy. Furthermore, the US Rice belt comprises five clusters, with economic importance varying from 13 to 45 %. These results will help us better allocate trials in advance and reduce costs without penalizing accuracy.

Pangenome-informed variant discovery to accelerate sorghum worldwide breeding programs

Brian R. Rice[1,4], John T. Lovell[2,3]*, Joanna Rifkin[2], Clara Cruet-Burgos[1] , Kerrie Barry[3], Adam Healey[2], Avinash Sreedasyam[2], Sujan Mamidi[2], Nadia ShakoorD, Todd MocklerD, Jane Grimwood[2], Jeremy Schmutz[2,3]*, Geoffrey P. Morris[1]

1 Department of Soil and Crop Science, Colorado State University, Fort Collins CO
2 Genome Sequencing Center, HudsonAlpha Institute for Biotechnology, Huntsville, AL
3 Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratories, Berkeley, CA
4 Department of Agronomy and Horticulture, University of Nebraska, Lincoln NE

The Sorghum genome is a pivotal resource for understanding drought tolerance in crops. To assist this effort, we present the most thorough and highest-quality exploration of sorghum diversity to date, including a pan-genome of 32 genetically diverse genotypes, deep re-sequencing of >1600 cultivars adapted to a wide range of environments, and complexity-reduction genotyping of >10,000 additional genotypes. We demonstrate how diversity within landraces facilitates the identification and characterization of causal variants. Population genetic analysis corrected for geography reveals substantial gene flow across the African continent and suggests that regional differentiation is largely due to local adaptation rather than allopatric divergence. However, extensive linkage disequilibrium hampers our ability to achieve gene-level resolutions for the contributions of distinct sequence variants. To address this limitation, we apply a SnpEff-enabled candidate gene approach tailored to identify novel variants crucial for drought tolerance. We focus on dhurrin, a cyanogenic glucoside unique to the Sorghum genus, and hypothesized to play a role in its adaptation to arid climates. Climate associations support the occurrence of climate selection on putative causal gene variants related to dhurrin biosynthesis and catabolism.

Platforms and methods for sharing and collaboration on agricultural genome to phenome using public and confidential data

Hao Cheng[1]*, Jack Dekkers[2], Juan Steibel[2], Gustavo de los Campos[3]

1 Department of Animal Science, University of California Davis, Davis, CA
2 Department of Animal Science, Iowa State University, Ames, IA
3 Department of Epidemiology and Biostatistics, Michigan State University, East Lansing, MI *qtlcheng@ucdavis.edu

Data sharing and collaboration are of increasing importance to enable validation, further research, and joint analysis of multiple data sets. However, these processes are often complicated or even prevented because public data may have limited visibility and accessibility, and private data often contain confidential or proprietary information, especially in the case of industry data.

With the long-term goal to advance Agricultural Genome to Phenome (AG2P) research and to increase rates of genetic improvement in agriculture, we are taking a tiered approach to fill the gap: 1) We foster sharing of public data when available; If this isn't feasible, we move to 2) innovating encrypted data sharing to protect privacy (see the abstract by Li et al. in this book of abstracts); Lastly, when sharing raw or encrypted data is impractical, we focus on 3) enabling collaborative research such as federated and transfer learning (see abstracts by Olabosoye et al., and de los Campos et al. in this book of abstracts). In addition to methodology development and evaluation, we are providing educational resources on effective data sharing and collaboration in AG2P, as well as associated statistical and data science skills, to various audiences, including undergraduates, graduates, plant and animal breeding professionals, and data-sharing stakeholders, such as field-based plant and animal breeders.

By facilitating access to and sharing of data for genomic and phenomic analyses, our hypothesis is that the availability of our multifaceted strategy will remove barriers that exist for the public sharing of research data and for the use of data and resources from industry in agricultural research. Removal of these barriers is essential to advance the field and to attain the long-term goal of out project. This research has been funded by AG2PI Seed Grant 2020-70412-32615, 2021-70412-35233, 2022-70412-38454 and USDA-NIFA AG2PI grant 2023-70412-41054.

Prospects for improving nondormant alfalfa using multi-omics approaches

Esteban F. Rios[1], Fernandes Filho[2], C.C., Sipowicz[3], P., Biswas, A.[1], Singh, A.[4], Messina, C.[5], Jarquin, D.[1]

1 University of Florida, Agronomy Department, Gainesville, FL
2 Centro de Tecnologia Canavieira, Piracicaba, Sao Paulo, Brazil
3 University of Florida, Plant Breeding Graduate Program, Gainesville, FL
4 University of Florida, Agricultural and Biological Engineering Department, Gainesville, FL

**Presenting and corresponding author: estebanrios@ufl.edu*

Alfalfa (Medicago sativa L.) is the most important perennial forage legume in the world because of its relatively high yield and nutritional value. Alfalfa is the fourth most valued crop in the USA with an estimated value of $9 billion. Warm and humid areas in the southeastern USA represent a large area for expansion of alfalfa as a forage crop. Alfalfa breeding for complex traits requires frequent and multiple phenotyping efforts within and across years, which is labor intensive and costly. Multi-omics data integration into plant breeding pipelines can result in greater genetic gain for complex traits. The objectives of this study were to explore strategies to improve predictive ability for complex traits in the University of Florida alfalfa breeding program: i) implement genomic prediction in alfalfa using new genotyping platforms, ii) deploy high-throughput phenotyping (HTP) in our alfalfa breeding program, iii) integrate environmental factors into prediction models to account for genotype by environment interactions (GxE), iv) perform phenomic prediction using spectral data from near infrared reflectance spectroscopy (NIRS) and unmanned aerial vehicles equipped with sensors. Genomic prediction has been applied in alfalfa for yield resulting in predictive abilities ranging between 0.2-0.4, and two genotyping platforms showed similar population structure and predictive ability. The inclusion of enviromics data in prediction models resulted in greater predictive ability for complex traits by accounting for GxE. As the capacity of classical field phenotyping is always limited, tools to assist selection have been developed. Remote sensing has enabled efficient and non-destructive estimation of yield and other traits in our alfalfa breeding program without compromising genetic gain. Our preliminary implementation of phenomic selection using near-infrared spectroscopy (NIRS) and hyperspectral sensors mounted to unmanned aerial vehicles resulted in greater predictive ability compared to genomic-based models for yield in most scenarios tested. Integrating genomics, enviromics, and phenomics could increase prediction ability of models and ultimately increase genetic gain for complex traits in alfalfa adapted to warm and humid agroecosystems in the southeastern USA.

Reduced representation bisulfite sequencing (RRBS) analysis shows that heat stress alters methylation of stress and metabolic genes in cattle

Rachel Reith[1], Renae Sieck[1], Pablo Grijalva[2], Rebecca M. Swanson[1], Anna M. Fuller[1], Duarte E. Diaz[2], Ty B. Schmidt[1], Dustin T. Yates[1], Jessica L. Petersen[1]

1 University of Nebraska – Lincoln, Department of Animal Science, Lincoln, NE
2 University of Arizona – Tucson, School of Animal and Comparative Biomedical Sciences, Tucson, AZ

Heat stress is a barrier for the beef industry due to its detriment to welfare and skeletal muscle growth. Heat stress regulates gene expression in part by altering DNA methylation. Understanding how heat stress changes the genomic structure may lead to new interventive strategies. Red Angus-crossed steers (n=6 per treatment) were assigned to thermoneutral (TN) or heat stress (HS) conditions (max temperature humidity index = 68, 85 respectively) for 21 days. Steers were pair-fed to those in HS. Muscle was biopsied from the longissimus dorsi on days 3, 10, and 21. Reduced representation bisulfite sequencing was performed on isolated DNA. Reads were mapped to ARS-UCD1.2, and methylation calls for CpG sites were extracted. Differential methylation analysis between groups for biopsy day revealed 563, 789, and 470 differentially methylated cytosines (DMC) in genes for the respective day, 58, 120, and 66 DMCs were in promoter regions (2000 bp 5' of a gene). Across all days, there were DMCs in the promoter regions of microRNAs, evidence of changes in gene regulation. DMCs in genes were input into DAVID to explore KEGG pathways. This revealed that heat stress affected 53, 57, and 26 pathways for the respective days. Several immune system pathways were affected on day 3, consistent with acute effects of heat stress. Heat stress affected PI3K-Akt and MAPK signaling pathways on all days, indicating changes in cell proliferation and survival. Changes in the HIF-1 pathway on day 3 suggest that muscle was responding to hypoxic-like conditions due to heat stress. This helps explain previously published RNA analyses from these animals, which showed oxidative stress in muscle due to heat stress. Overall, these data show that heat stress altered methylation of genes related to cell signaling, stress, and metabolism, all of which can be functionally impaired in heat-stressed animals.

Selecting single nucleotide polymorphisms (SNPs) to implement a multi-trait genomic selection pipeline for root quality and the cassava leaf spot disease complex

Gabriel Mamedio de Freitas[1], Diana Solarte[1], Jean-Luc Jannink[3], Eder de Oliveira[2], Antonio Garcia[1]

1 College of Agriculture "Luiz de Queiroz" (ESALQ), University of São Paulo (USP), Genetics Department, Piracicaba, São Paulo, Brazil
2 EMBRAPA (Brazilian Agricultural Research Corporation), Cruz das Almas, Bahia, Brazil
3 United States Department of Agriculture, Agricultural Research Service (USDA-ARS), Robert W. Holley Center, Ithaca, NY

Genomic selection has promoted significant advances over traditional methods that depend exclusively on phenotypic selection. Given the specific challenges in cassava (Manihot esculenta) production, such as susceptibility to leaf diseases and the search for increased productivity, this approach has proved fundamental for the genetic breeding of the crop. In this context, the strategy has resulted in the provision of clones better adapted to the adverse conditions found in the Americas and on the African continent. This study aimed to apply multitrait genomic selection to evaluate the severity of the leaf disease complex and the production of dry roots in cassava accessions from the active germplasm bank of Embrapa Mandioca e Fruticultura, Brazil. Cassava clones at different stages of the selection cycle were genotyped and phenotyped. Genotyping was done using DArTseq, and the genomic data was aligned with the reference genome v8.1 using BWA. SNP identification was conducted using the Genome Analysis Toolkit (GATK), followed by genotype data selection using the TASSEL software, with a Minor Allele Frequency (MAF) of 1% and a minimum call rate of 80%, followed by imputation. In all, 20,625 informative SNPs were retained. Phenotyping was carried out between 2017 and 2019, covering the evaluation of the severity of symptoms of the foliar disease complex caused by Passalora vicosae, Clarohilum henningsii, and Passalora manihotis, as well as the dry root yield (DRY). The latter was calculated by multiplying T.FRY (Total fresh root yield) and DMC (Dry matter content). Subsequent steps will include obtaining BLUE estimates and implementing the relevant multi-trait genomic selection models. This study is not limited to providing an extensive analysis of cassava's genomic and phenotypic characteristics but also aims to contribute to developing more resilient and productive varieties. The impacts of this work extend to industrial use and human consumption of cassava in Brazil.

Testing theories of genetic variation in complex traits using machine learning techniques

Kira Villiers[1], Helen Huang[2], Zhi Chen[2], Ben J. Hayes[1]

1 Queensland Alliance for Agriculture and Food Innovation (QAAFI), The University of Queensland, St. Lucia, Queensland, Australia
2 School of Electrical Engineering and Computer Science, The University of Queensland, St. Lucia, Queensland, Australia

Agricultural breeders depend on the existence of genetic variation to achieve year-on-year genetic gains. The theory of how genetic variation is established, maintained, and lost is not settled science in quantitative genetics. A large barrier to comparing different proposed theories and their predictions is the theories' dependence on hard-to-measure parameters, like strength of stabilising selection and magnitude of the effects of new mutations on the trait. We investigated the potential of simulating population genetic histories according to theories of genetic variation, and using these to train machine learning models for predicting the values of parameters of genetic variation for real datasets. If this could be achieved, it would unlock more avenues for investigation and comparison of the theories of genetic variation, which would allow breeders to make informed decisions about their management of genetic diversity. Simulated histories corresponding to the breeding program design of the University of Illinois' Long Term Selection Experiment were produced under different values of 10 parameters of genetic variation. Of the search methods and machine learning methods tested, the long short-term memory (LSTM) recurrent neural network model produced the highest accuracies of prediction of variation parameters on simulation datasets. However, the parameters the LSTM predicts for a real dataset are parameters that, in simulation, produce trends that are divergent from the behaviour observed in the real dataset. The low number of dimensional features, and hence low power, of the agricultural-style simulated datasets seem a likely cause of this gap. The use of several varied datasets or higher-detail datasets could be measures that would extend this method into a useful approach for predicting historical parameters of genetic variation.

The African catfish (Clarias gariepinus) reference genome: A near telomere-to-telomere (T2T) de novo chromosome-level assembly

Tom Goldammer[1,2], Julien A. Nguinkal[1,3], Ronald M. Brunner[1], Yedomon A.B. Zoclanclounon[4], Yutang Chen[5]

1 Research Institute for Farm Animal Biology (FBN), Fish Genetics Unit, Dummerstorf, Germany
2 University of Rostock, Faculty of Agriculture and Environmental Sciences, Molecular Biology and Genetics of Fish, Rostock, Germany
3 Bernhard-Nocht Institute for Tropical Medicine, Department of Infectious Disease Epidemiology, Hamburg, Germany
4 Jeonbuk National University, Department of Crop Science and Biotechnology, Jeonju, South Korea

The African catfish is an ecologically and commercially important freshwater air-breathing catfish that can withstand numerous environmental conditions and farming practices. Its high fillet quality is similar to wild salmonids. Therefore, considerable efforts are being made to increase aquaculture and optimize harvest yields of the African catfish, one of the most commonly farmed clariids. We sequenced the genome of the African catfish and provide by this one prerequisite for the development of genetic breeding tools. We generated a gapless telomere-to-telomere (T2T) de novo chromosome-level assembly with high-resolution haplotypes, by integrating long-range sequencing (Hi-C) with PacBio single-molecule (HiFi), Oxford Nanopore, and Illumina sequencing data. The diploid genome assembly yielded 58 contigs with a total length of 969.7 Mb and a contig N50 of 33.7 Mb. We reported 25,655 predicted protein-coding genes and 49.94% repetitive elements in the C. gariepinus genome. Interspersed repeats are the most abundant class of repetitive elements (46%). Retroelements and DNA transposons accounted for only 12 and 6 percent of the repeatome, respectively. Approximately 99% of the assembled genome is spanned by the 28 chromosomes of the Primary assembly, which have no gaps. The comparative phylogenetic analyses performed with catfishes and two outgroup species (common carp and goldfish) identified the split of air breathing catfishes (Claridae clade) as a monophyletic group around 98 Mya, which is roughly comparable to the divergence time between rodents and humans (96 Mya). Our genome assembly provides the first comprehensive gene annotation and haplotype information, such as the male-specific haplotype, enabling us to identify critical genes and molecular mechanisms underlying amphibious traits and terrestrial adaptation of air breathing catfishes. In sum, the data will serve as a valuable resource for future studies in elucidating these unique biological traits in related teleost's and leverage these insights for aquaculture improvement. The project was funded by the EMFF under the number MV-II.1-LM-014.

The utility of landrace environmental data for climate-adaptive maize breeding

Forrest Li[1], Daniel Gates[2], Edward Buckler[3], Sarah Hearne[4], Daniel Runcie[2], Jeffrey Ross-Ibarra[5]

1 Department of Evolution and Ecology and Department of Plant Sciences, University of California Davis, Davis, CA
2 Department of Plant Sciences, University of California Davis, Davis, CA
3 United States Department of Agriculture, Agricultural Research Service (USDA-ARS), Robert W. Holley Center for Agriculture and Health, Ithaca, NY, and Institute for Genomic Diversity, Cornell University, Ithaca, NY
4 International Maize and Wheat Improvement Center (CIMMYT), El Batan, Texcoco, Estado de Mexico, Mexico
5 Department of Evolution and Ecology, Center for Population Biology, and Genome Center, University of California Davis, Davis, CA

Maintaining crop production yields in the face of climate change is a major challenge facing plant breeding today. Considerable adaptive genetic variation to improve breeding programs exists in ex situ germplasm collections of traditional landraces, but effective utilization of this diversity requires identifying adaptive alleles. Furthermore, testing the agronomic performance of large populations of diverse germplasm in field trials is costly, and best practices to identify favorable alleles or select varieties in landrace collections are still in development.

To evaluate approaches for identifying environmentally adaptive alleles, we collected climatic variables from the environment-of-origin of each of nearly 4000 previously genotyped traditional maize landraces representing a breadth of maize diversity. We evaluated three different approaches of using climate data to identify useful diversity based on success in predicting phenotypic performance in field trials. First, we performed gene-environment association (GEA) studies identifying genetic loci associated with environmental heterogeneity. We identify loci such as hsftf9, a putative heat shock protein that may provide adaptive benefit in high temperatures, as well as large-scale inversions such as Inv4m, previously described to be locally adaptive to highland conditions. Second, we used existing genotype data to predict optimal climate niches that may inform selection of landraces better adapted to target environments as well as demonstrate the limitations of prediction of environmental variables across spatial ranges. Third, we tested the ability of genomic prediction using population structure and high-resolution SNP data for landrace collections.

We find that genomic prediction capturing population structure and genetic relationship continues to have the highest predictive breeding value accuracy compared to models employing climate-of-origin variables or models focused on large-effect loci associated with climate.

Toward a cattle and sheep pan-epigenome: Elucidating the functional impact of DNA and RNA methylation on gene expression

Shangqian Xie[1], Morgan R. Stegemiller[1], Gabrielle M. Becker[1], Kimberly M. Davenport[2], Katie A. Shira[1], Denise Konetchy[1], Patricia Villamediana[3], Brenda M. Murdoch[1], Stephanie D. McKay[4]

1 University of Idaho, Department of Animal, Veterinary, and Food Sciences, Moscow, ID
2 Washington State University, Department of Animal Sciences, Pullman, WA
3 South Dakota State University, Department of Dairy and Food Science, Watertown, SD
4 University of Missouri, Division of Animal Sciences, Columbia, MO

Elucidating epigenetic impacts on phenotypic variation involves characterization of multiple epigenetic modifications including DNA 5-methylcytosine (5mC) and RNA N6-methyladenosine (m6A). These modifications play multifaceted and pivotal roles in biological processes, such as gene silencing, activation, cell differentiation, disease and adaptation of organisms. While previous studies have examined the role of DNA 5mC and RNA m6A methylation in regulating gene expression, the intricate relationship among DNA methylation, RNA methylation and gene expression remain to be fully elucidated. In this study, we collected DNA 5mC methylation, RNA m6A methylation and RNA-seq data from 24 samples consisting of three tissues (caruncle, mammary and spleen) from four cattle (N = 12) and sheep (N = 12) and proposed a statistical model to quantify the effects of DNA and RNA methylation on gene expression. The results indicated that (1) RNA methylation has a greater influence on gene expression compared to DNA methylation in gene-body and promoter regions; (2) genes containing DNA methylation within their bodies are more likely to exhibit RNA methylation in their gene-bodies as opposed to the promoter region of these genes; (3) genes with DNA methylation in their promoter regions are less likely to have RNA methylation in their promoter or gene-body; (4) the expression ratio of genes that have both DNA and RNA methylation in their gene-body are higher than those in genes with both DNA and RNA methylation within promoter regions; (5) the effects of DNA and RNA methylation in a gene-body are positively correlated with regulation of gene expression. These findings enhance our understanding of genome-wide dynamics and the complex associations among DNA, RNA methylation and gene expression, offering new perspectives and opportunities to further animal epigenetics research. Notably, this work supports the establishment of a cattle and sheep pan-epigenome, contributing to a better understanding of how different genomes, epigenomes and gene products from diverse breeds affect a variety of important biological phenotypes, supporting more accurate prediction of traits and improving breeding strategies.

Using encrypted genotypes and phenotypes for reproducible and collaborative genomic analyses to maintain data confidentiality

Donna Li[1], Tianjing Zhao[2], Jack Dekkers[3], Richard Mott[4], Hao Cheng[1]

1 University of California, Davis, Department of Animal Science, Davis, CA
2 University of Nebraska-Lincoln, Department of Animal Science, Lincoln, NE
3 Iowa State University, Department of Animal Science, Ames, IA
4 University College London, Department of Genetics, London, United Kingdom

As the cost to genotype individuals has decreased, the use of genotypic and phenotypic data has become increasingly popular for genome-to-phenome analysis. In the case of collaborative studies and reproducible research, data owners may prefer or are required to keep their data private due to privacy concerns, commercial interest, or data sharing policies. Past research has considered different methods to protect genetic information, such as deidentification and/or encryption. However, most of these strategies have their shortcomings as information about individuals may still be inferred from unencrypted genotypes, and encryption schemes often require genomic data to be decrypted before analysis. Homomorphic encryption, however, allows specific computations to be performed on encrypted data and provides the same results, decrypted if necessary, as those obtained from the same computations on unencrypted data. In this study, we extended a method called homomorphic encryption for genotypes and phenotypes (HEGP), which applies homomorphic encryption to genomic and phenomic data, to wider applications. We show that HEGP can be effectively applied to many popular mixed models, beyond single-marker regression. These models, including Bayesian variable selection methods such as those in Bayesian Alphabet, are routinely employed in the fields of animal and crop improvement, for genetic analyses of quantitative traits, including genetic parameter estimation, genomic prediction, and GWAS. By testing HEGP on both simulated and real data in pigs, we evaluate the correlation between estimates of marker effects and breeding values based on the encrypted and unencrypted data. In addition, we also compare GWAS results from encrypted and unencrypted data, using Bayesian variable selection methods. We demonstrate that HEGP successfully obscures information about individuals for data privacy, while generating identical estimates of marker effects using the encrypted and unencrypted data, without the need for decryption. This research has been funded by AG2PI Seed Grant 2020-70412-32615 and 2021-70412-35233, and USDA-NIFA AG2PI grant 2023-70412-41054.

Applying the Science - Scaling Science Innovation for Use

A high-throughput method for the investigation of tick, pathogen, and commensal microbiome associations at single-tick resolution

Rachel F. Daniels[1,2], Gaurav Chauhan[1], Patrick W. Marsh[3], Jeffrey A. Bailey[3], Diane P. Genereux[2], Elinor K. Karlsson[1,2]

1 Genomics and Computational Biology, University of Massachusetts Chan Medical School, Worcester MA
2 Vertebrate Genomics, Broad Institute of MIT and Harvard, Cambridge, MA
3 Department of Pathology and Laboratory Medicine, Warren Alpert Medical School, Brown University, Providence, RI

Global climate change has led to an increasing prevalence of tickborne diseases, with direct effects on both livestock and outdoor workers. The development of effective control strategies requires improved understanding of the transmission dynamics of these pathogens, including the complex interactions between the tick microbiome and co-infection with other pathogens. However, these investigations require large-scale, high-resolution data. Here we present a method that combines citizen science with new molecular strategies to provide data at single-tick resolution to inform the prevention and management of tickborne pathogens.

We implemented a citizen-science approach to collect >3,000 ticks from across the continental US and developed high-throughput targeted sequencing methods to identify tick and host species, as well as viral and bacterial pathogens associated with tickborne infections, from individual ticks.

Molecular analysis identified seven tick species, including Amblyomma americanum (Lone star tick), Dermacentor andersoni (Rocky mountain wood tick), Dermacentor variabilis (American dog tick), Haemaphysalis longicornis (Asian longhorned tick), Ixodes pacificus (Western blacklegged tick), Ixodes scapularis (Blacklegged tick), and Rhipicephalus sanguineus (Brown dog tick). Among the mammalian host species, we detected ticks that carried host DNA from humans, dogs, and cattle, with one tick containing both human and dog DNA.

In our pilot analysis of the microbiomes of nearly 200 ticks, more than half (61.2%) were infected with Borrelia burgdorferi, the causative agent for Lyme disease. Our findings also validated previous findings of overall higher microbial diversity in males than in females and the higher abundance of Rickettsia in the female ticks. We also detected 13 instances of single ticks infected with multiple pathogens and identified bacterial species pathogenic to humans, including four not previously reported in association with ticks, thus illustrating the potential of these methods for both surveillance and detection of emerging pathogens.

Our results demonstrate the power of high-throughput sequencing at single-tick resolution to improve our understanding of the complex system of transmission dynamics in tickborne disease, which could lead to new strategies for pathogen control.

Allele mining to recover key drought and striga tolerant alleles from sorghum pangenomes

Samuel Abebrese[1*], Clara Cruet-Burgos[1], S M Abdul-Awal[1], Joel Masanga[2], Geoffrey Morris[1]

1 Colorado State University, Department of Soil and Crop Science, Fort Collins, CO
2 Penn State University, Eberly College of Science, Department of Biology, Pennsylvania, PA

*samuel.abebrese@colostate.edu

Pangenomes are valuable resources for mining alleles of genes of relevant agronomic traits for crop improvement. However, testing hypotheses on causative variants can be slow and time consuming. Here, we summarize a new pipeline in sorghum, the global food security crop, that identifies putative causative variants for drought and striga tolerance and enables their rapid introgression to test hypotheses on causative variants. Comparative genomics and environmental genome-wide association studies (eGWAS) identified putative causative variants in several genes, such as DROUGHT1 (SbDROT1), LOW GERMINATION STIMULANT 1 (LGS1) and PHENYLALANINE/TYROSINE AMMONIA-LYASE (PTAL) which may underlie variation in drought and striga tolerance. Allele-specific markers for putative causative variants are designed through Kompetitive Allele-specific PCR (KASP) technology to screen new backcross introgression families and guide the development of Near Isogenic Lines (NILs) through foreground selection. Leveraging Mendelian sampling variance, mid density genome-wide markers are employed to identify NILs with higher recovery of the recurrent parent genome in early backcross generations to test hypotheses on causative variants. This pipeline should accelerate delivery of valuable climate resilience alleles harbored in sorghum pangenomes to breeding programs for the benefit of farmers in climate impacted regions of Africa.

Everything everywhere all at once: Unveiling microbial dynamics during forest-to-oil palm plantation soil transition in Sarawakian soils

Zahidah Ayob[1], Nor Azizah Kusai[1], Nur Maisarah Jantan[1], Meilina Ong Abdullah[1]

1 Malaysian Palm Oil Board, Bandar Baru Bangi, Selangor, Malaysia

Transformation of forest to crop land is believed to significantly impact the soil microbial community. While previous studies have aimed to characterise the microbial shifts of land conversion to oil palm plantations via amplicon sequencing, knowledge of the specific roles played by individual microbes still remains elusive. The purpose of this investigation is to generate metagenome-assembled genomes (MAGs) using shotgun metagenome sequencing with genome binning for a more in-depth analysis from three distinct soil types: secondary forest (SF), 3-year-old oil palm plantation (SAJ), and 11-year-old oil palm plantation (SAB) in Sarawak, Malaysia. Through individual de novo metagenome assembly of 12 soil samples and a self-supervised contrastive learning approach for genome binning, coupled with genome dereplication, we successfully recovered 462 MAGs exhibiting over 50% genome completeness and less than 10% contamination. Taxonomic assignment based on the latest Genome Taxonomy Database (GTDB) revealed the presence of MAGs spanning 23 phyla, 30 classes, 59 orders, 79 families, and 157 genera. Intriguingly, only four MAGs could be classified at the species level, indicating that the majority of MAGs recovered from Sarawakian soil represent novel species absent from the GTDB. Sequencing read mapping to the recovered MAGs demonstrated that, on average, 60% (ranging from 50% to 68%) of reads could be attributed to these MAGs. Notably, MAGs affiliated with the genus RAAP2 exhibited the highest average relative abundance among the soil samples, followed by the genus Terracidophilus. Both genera are commonly associated with acidic conditions, and this finding is consistent with their isolation source (acidic soil). In the secondary forest, there is a pronounced prevalence of members from the family Isopharaeceae (chemoorganotroph) and UBA183. Conversely, in the 3-year-old and 11-year-old plantations, members of the family Syntrophobacteraece (sulfate reducer) and Palsa_688 exhibit a significant increase in relative abundance, respectively. Interestingly, both oil palm plantations exhibit a comparable heightened relative abundance the Bog793 family members, of which are starkly absent or rarely observed in the secondary forest soil. These findings shed light on the intricate microbial dynamics in soil subjected to land conversion, with potential implications for sustainable land use practice.

High-density phenotyping and genomics to improve feed efficiency in pigs

Hélène Gilbert[1], Florence Gondret[2], Vanille Déru[3], Céline Carillier-Jacquin[1], Juliette Riquet[1]

1 GenPhySE, INRAE, Castanet-Tolosan, France
2 PEGASE, INRAE, Institut Agro, Saint-Gilles, France
3 INP-Purpan, Université de Toulouse, Toulouse, France

Improving feed efficiency stands as crucial for enhancing livestock sustainability, as it contributes to minimizing production costs animal excretion, and gas emissions due to feed production. Achieving superior feed efficiency necessitates the recording of individual feed intake, a challenge overcome in pigs through costly automatic feeders and RFID transponder-equipped animals. Alternatively, indicators of feed efficiency can be searched in accessible samples like blood and feces. In blood sampled at different periods, transcriptomic information has proven valuable, achieving high accuracy in predicting feed efficiency classes (100%) and quantitative measures (R^2=0.80) through machine learning. In both cases, combinations of around 30 gene transcripts were retained as reliable predictors. Combinations between blood molecular networks with the blood metabolome information also provided valuable information on the biology behind feed efficiency phenotype. Feces also provide accessible information: near-infrared spectroscopy was used to develop predictors of individual digestive coefficients, indicative of the efficiency in the first stage of energy and nutrient utilization by animals. Fecal sampling further allows approximation of individual gut microbiota composition through partial sequencing (16S) or full metagenome sequencing. Variability in the gut microbiota composition has then emerged as a potential predictor of digestive and feed efficiency. However, determining the genetic architecture of feed efficiency, a composite trait resulting from multiple biological functions, is still challenging. Yet, with the advent of complete genome sequencing, combining imputed sequence variants and phenotyping in two connected pedigreed pig populations has yielded a substantial dataset- (17,033,057 SNP imputed in 3,505 pigs), facilitating the detection of novel genomic regions influencing feed efficiency, related production traits and microbiota traits. Altogether, these high-throughput techniques provide an improved understanding of the factors influencing feed efficiency, paving the way for targeted improvements of livestock sustainability.

Leveraging canine citizen science for cancer genomics research: Development of genetic screening and genomic selection tools in dogs

Frances Chen[1,2], Kate Megquier[2], Shauneen O'Neill[3], Brittney Kenney[1,2], Madeline Zimmermann[4], Jane Russenberger[4], Eldin Leighton[4], Breno Fragomeni[3], Elinor Karlsson[1,2]

1 UMass Chan Medical School, Genomics and Computational Biology; Molecular Medicine, Worcester, MA
2 Broad Institute of MIT & Harvard, Vertebrate Genomics, Cambridge, MA
3 University of Connecticut, Animal Science, Hartford, CT
4 International Working Dog Registry, San Antonio, TX

Dogs serve diverse roles in society, from working in agriculture, security, and the military, as guide and assistance dogs, and as beloved family members. They are also a valuable comparative model for cancer genomic research due to key similarities with humans, including spontaneously acquired disease, shared tumor mutations, similar environmental exposures, and comparable treatment modalities. Importantly, their high incidence rates, shorter lifespans and accelerated disease course allow for more feasible longitudinal cancer outcome studies.

Applying genomic technologies in canine cancer research presents a unique opportunity to enhance our understanding of cancer biology, benefiting both canine and human health. It also holds promise for directly reducing cancer incidence in the dog population through genetic selection.

Currently, we are utilizing genetic and longitudinal phenotypic datasets from thousands of individuals in the Golden Retriever Lifetime Study (GRLS) and the International Working Dog Registry (IWDR) to not only investigate genetic contributions to canine hemangiosarcoma, but also develop genomic tools for directly reducing cancer incidence in dogs. Our group has previously identified heritable risk variants for hemangiosarcoma in specific breed populations and are now leveraging these cohorts for larger genome-wide association studies (GWAS) to uncover known and novel cancer risk variants. Simultaneously, we will employ these same data to generate genomic estimated breeding values (GEBVs). Building on our success in using estimated breeding values to reduce mast cell cancer in guide dogs, we will collaborate with our working dog community to operationalize GEBVs and demonstrate proof of concept of genomic selection. Disease-associated variants will refine GEBVs and contribute to improving genetic screening and genomic resources for the entire canine community. Finally, our canine community science platform at Darwin's Ark will facilitate testing and refinement of these tools while further driving discovery in comparative oncology.

(continue to page 60)

(continued from page 59)

In summary, our work aims to leverage genetic and longitudinal data from extensive canine cohorts to investigate hemangiosarcoma, identifying heritable risk variants and developing genomic tools for cancer reduction. Through collaborative efforts and the integration of genomic approaches, we aim to advance genetic selection methods in dogs, ultimately aspiring to reduce hemangiosarcoma incidence in the broader canine population.

MARPLE diagnostics: Real-time strain identification of wheat rust fungi

Loizos Savva[1], Oscar Gonzalez[1], Diane Saunders[1]

1 John Innes Centre, Norwich Research Park, Norwich, United Kingdom

Wheat rusts have been associated with crop failures and famine throughout history and recent changes in the pathogen populations have expanded their ability to infect different wheat varieties, causing a serious threat to global wheat production. To address this, we developed the MARPLE diagnostics tool to improve the speed and resolution of current surveillance and diagnostic systems.1 This tool makes use of a diverse selection of wheat rust pathogen isolates, sourced from around the world, whose genetic variations have been characterised and used to identify key genetic markers associated with the different pathogen strains. This foundational groundwork permits the classification of field isolates into distinct genetic lineages. Utilizing the nanopore sequencer it enables rapid identification of both newly emergent strains and those with specific properties such as fungicide resistance, generating results within 48 hours of field sampling. MARPLE achieves near real-time strain identification of wheat rust fungi, equipping researchers with a tool for proactive disease management.

Modified Tip Surface (MTS) is a high-quality, high-speed, and high-throughput DNA extraction alternative to magnetic bead purification, yielding assay-ready DNA without using alcohol or high salt conditions.

Adam Majot[1], Wilco de Jager[2], Pawan Basnet[1], Michael Dercoli[1], Daniel Carrera-Lopez[1], Niels Kruize[2]

1 Agriplex Genomics, Cleveland, OH
2 Agriplex Genomics, Veenendaal, Utrecht, Netherlands

High throughput genotyping programs rely on the efficiency of DNA extraction, next-gen sequencing library preparation, and data processing. Both library preparation and data processing options have quickly evolved to cater to a wide variety of formats and price points. However, DNA extraction technology has largely remained static. High prices are coupled with long prep times and, therefore, slower throughput. In return, customers get purified DNA that is suitable for most downstream applications. Alternately, lower-priced methods such as alkaline lysis are amenable to cost considerations and are ultra-high-throughput-friendly – the tradeoff with such methods is that DNA quality is not suitable for the same range of applications as with purified DNA.

The Modified Tip Surface (MTS) extraction platform offers a leap forward in extraction capabilities, combining all the necessary requirements of ultra-high-throughput workflows (reduced cost, increased speed, alcohol-free) with an output of highly purified DNA that is suitable for a full suite of genomics/NGS workflows. First, we compared yield on lysis buffer composition (chaotrope+/-). Regardless of DNA extraction platform (comparing two paramagnetic-bead based products and MTS), chaotrope-free lysis buffer resulted in greater DNA yield, as assessed by absorbance and rtPCR, than that from chaotrope+ buffer when run on nasal swab samples. Additionally, nasal swab extraction yield was not improved by the inclusion of alcohol wash steps, which is of critical importance to the simplification of workflows on automated decks where alcohol condensation often poses significant challenges during high-throughput processes.

Swab extracts were functionally tested for quality using the Agriplex multiplex PCR Bovine Parentage genotyping panel. The Agriplex Bovine Parentage panel interrogates 552 of 554 ICAR-recommended parentage SNPs and includes all 200 ISAG-recommended parentage SNPs. MTS extract performance was optimized by titrating DNA input concentration against performance metrics (NGS read recovery and call rate) and compared head-to-head against competitor-derived purified DNA and crude ear-notch lysate. MTS offers an alternative, more affordable approach to DNA extraction that features value and efficiency for most downstream NGS applications and facilitates a speed that is convenient to a range of operational sizes, from small scale to the most high-throughput genomics lab operations.

NEBNext® streamlined solutions for plant and animal genomics: Novel library prep workflows and automated genotyping solutions

Kruti Patel[1], Anetta Nowosielska-Vecchietti[1], Michael Petridis[1], Karen McKay[1], Chelsea Pinegar[1], Kaylinette Pinet[1], Chen Song[1], Pingfang Liu[1], Chaithanya Ponnaluri[1], Gautam Naishadham[1], Matthew Campbell[1], Brad Langhorst[1], Christine Summer[1], Brendan Desmond[1], Bjoern Textor[1], Andrew Barry[1], Keerthana Krishnan[1]

1 New England Biolabs, Ipswich, MA

Next generation sequencing enables the rapid interrogation of both DNA and RNA allowing for genome analysis, gene expression detection, and variant or marker identification using whole genome or target enrichment panels. As sequencing technologies improve and capacities expand, there has been a need to improve library construction methods for speed, scalability, and ease of use without compromising library or data quality. Here, we describe methods for streamlined and automatable library preparation for whole genome, transcriptome, and target enrichment sequencing analysis of various plant and animal species.

For RNA and DNA workflows, we have developed new protocols (NEBNext UltraExpress DNA, NEBNext UltraExpress FS DNA, and NEBNext UltraExpress RNA protocols), which enable users to generate robust sequencing libraries quickly and easily from a variety of sample types. The option of a single protocol for all input amounts within the kit range improves ease of use and supports automation. These protocols implement master-mixed components, shorter incubation times, and fewer bead clean-up steps to significantly decrease sample-handling and streamline the process. We demonstrate the robustness of these kits by evaluating uniformity of coverage across the genome (DNA) or consistent 5'-3' transcript coverage and transcript abundance (RNA) across a range of inputs and sample types, including Human, Arabidopsis, Rice, Maize, and bacterial samples.

Additionally, to meet the need for high-throughput screening applications, such as marker assisted selection, we have developed a fully automated protocol for target enrichment using our NEBNext Direct Genotyping Solution. This technology facilitates high-throughput sample processing including pre-capture barcoding and pooling of up to 96 samples per capture. As genotyping applications require workflows to be cost-effective and scalable, we incorporated automation to both miniaturize reaction volumes and prepare libraries. This process enables a fully automated workflow and reduces the operational costs per sample while preserving data quality. Sequencing parameters of uniformity, enrichment specificity , and mean target coverage were used to evaluate the performance of tomato DNA enriched with a 500 marker panel.

The protocols described here enable fast, cost-effective, and scalable library preparation methods for generating high-quality libraries and sequencing data relevant for both plant and animal species.

Next-generation plant breeding: Harnessing pangenomics to drive gene discovery and selection in US and international breeding programs

Josh Clevenger[1], Walid Korani[1], Patricia Sanmartin[1], Abigail Burch[1], Kelly Chamberlin[2], Jorge Baldessari[3], Peggy Ozias-Akins[4], Corley Holbrook[5]

1 HudsonAlpha Institute for Biotechnology, Huntsville, AL
2 United States Department of Agriculture, Agricultural Research Service (USDA-ARS), Stillwater, OK
3 INTA Estación Experimental Agropecuaria Manfredi
4 University of Georgia, Tifton, GA
5 United States Department of Agriculture, Agricultural Research Service (USDA-ARS), Tifton, GA

The use of pangenomes in applied plant breeding is in its infancy, yet we need tools now to drive crop improvement for long-term yield gain and short-term reaction to new threats. The advent of third-generation sequencing allows us to effectively leverage pangenomes in applied breeding programs of any size. We showcase the implementation of pangenomics across continents for gene discovery and accelerated selection to drive crop improvement for disease resistance. Peanut smut is a fungal pathogen in Argentina that significantly affects global supply chains. Using pangenomics, we discovered functional variation linked to smut resistance and selected smut-resistant lines that retain high levels of agronomically elite background. We immediately began developing resistant varieties for Argentina and breeding for smut resistance in germplasm suitable for the United States. The power of effectively deployed pangenomics allowed us to breed resistant varieties ready to deploy in the United States if and when peanut smut moves to North America.

Simplified workflow for faster genotyping by sequencing using partial combinatorial dual barcodes and an optimized analysis pipeline

Jennifer Lavoie[1], Angela Burrell[1], Krishna Reddy Gujjula[1]

1 Thermo Fisher Scientific, Austin, TX

Targeted Genotyping by Sequencing (GBS) is a robust and cost-effective method for marker-assisted breeding and selection. Targeted GBS provides a scalable workflow, allowing for thousands of markers to be genotyped in up to thousands of samples per day. As GBS use expands, demand increases for higher throughput, faster turn-around times, and increased multiplexing capability.

To achieve higher throughput with less material and greater ease of use, we propose a new library prep workflow using partial combinatorial dual barcodes for use with the Ion GeneStudio™ S5 System that not only increases available barcodes, but also achieves complete, normalized libraries for up to 3072 samples in less than a day. Additionally, we propose an analysis workflow optimized for more than 6000 samples to be sequenced and genotyped in less than 24 hours using a single Ion GeneStudio™ S5 System. These optimizations enable completion of genotyping of >30,000 samples within a 5-day week. Here we report early testing results showing high quality results achieved using this novel workflow and analysis pipeline.

For Research Use Only. Not for use in diagnostic procedures.

Skim sequencing as an affordable alternative for GWAS in rice

Maria Ymber Reveche[1], Dmytro Chebotarov[1], Nobe Atanacio[1], Alice Laborte[2], John Damien Platten[1]

1 International Rice Research Institute, Rice Breeding Innovations Department, Los Baños, Laguna, Philippines

2 International Rice Research Institute, Sustainable Impact through Rice-based Systems Department, Los Baños, Laguna, Philippines

Applications requiring mid- to high-density genotype data have typically faced a tradeoff: increasing marker density is closely correlated with increasing price. Whole-genome sequencing is prohibitively costly per sample to allow any study involving adequate numbers of entries. Consequently applications requiring mid-density data have tended to focus on complexity-reduction techniques such as oligo microarrays, amplicon sequencing, GBS and target capture to reduce the genomic material to be assayed.

In this study, skim-sequencing was used as a cost-effective alternative to GBS and other techniques for genotyping rice landraces in an environmental GWAS study. Over 6500 rice accessions were selected based on availability of high-confidence location of origin data. These accessions were sequenced to a depth of 2-3 fold (average 2.5x) at the Texas A&M AgriLife Genomics and Bioinformatics Service facility at an average cost of 30USD per sample plus approximately 3USD for DNA extraction through Intertek. SNP calling from the first 1400 accessions revealed an average of approximately 3M filtered SNPs from Oryza glaberrima and 2M from O. sativa accessions. As expected for such low-coverage data, the overlap of SNP sites between samples was relatively poor; only around 9000 sites had a high call rate across 308 O. glaberrima accessions. However, imputation using standard pipelines was able to improve this dataset to >0.5 M SNPs.

Application of this imputed dataset to environmental variables revealed putative loci contributing significant associations with high day and night temperatures, and low relative humidity. The putative favourable haplotypes originated from stress-prone regions of central Mali, on the Niger River skirting the edge of the Sahara region. Validation of these associated regions is now underway.

Skim-sequencing is now quite competitive with complexity-reduction techniques, with costs comparable to GBS, and substantially less than high-density arrays (up to $100/sample). It also offers significant benefits: it makes fewer assumptions about the starting material and is far easier to generalize results across studies, making these results a substantial resource for future studies.

Informatics and Data Analytics

A systems biology approach in profiling the Sorghum bicolor transcriptome toward the development of routine breeding toolkits

Susan Moenga[1], Damaris A. Odeny[2]

1 International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Nairobi, Kenya
2 Center of Excellence in Genomics and Systems Biology, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Crop Physiology and Modelling Cluster, Hyderabad, Telangana State, India

Sorghum (Sorghum bicolor (L.) Moench) is the fifth most important cereal crop globally with versatile end uses including food, beverage, feed, forage, and fuel. As a C4 plant with more efficient photosynthesis under dry and hot environments, sorghum is an ideal candidate for sustainable livelihoods and has been promoted as climate resilient with the exceptional ability to thrive under low-input conditions. Its relatively small genome size further makes it an attractive model for studying tropical cereal functional genomics. Consequently, significant molecular data has been generated over the years for sorghum with the goal of deeper understanding of the molecular mechanisms underlying its versatility. The objective of the current study was to utilize the publicly available sorghum transcriptome datasets to better understand photoperiod sensitivity ranges across different end-use sorghum types. We analyzed a large collection of published transcriptomes of six sorghum genotypes representing different end-uses and photoperiod sensitivity within prescribed developmental contexts. The genotypes included were Tx430 (grain sorghum), Wray (sweet sorghum), R.07007 and TX08001 (biomass sorghum), as well as 100M and 58M both of which are representative of the photoperiod sensitivity spectrum. Expression and co-expression network modules characterized in this study will provide insights into spatially, genetically, and developmentally resolved programs and the regulatory circuits associated with trait and subsequent use divergence in sorghum. Candidate genes (and modules) for different use cases as well as niche environmental adaptation will be predicted by use of machine learning approaches, expanding the available toolkit for sorghum breeding. Structural variations identified across contrasting phenotypes will be converted into quick routine assays for future characterization and improvement of tropical sorghum genotypes for their response to photoperiod sensitivity. The outcome will also inform future studies on the effective utilization of existing public datasets for practical development of routine breeding tools.

Conservation of gene regulatory elements in plants

Abraham Morales[1], Leo Baumgart[1], Sharon Greenblum[1], Peng Wang[1], Yu Zhang[1], Ronan O'Malley[1,2]

1 U.S. Department of Energy, Joint Genome Institute, Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA
2 Department of Human Genetics, University of Chicago, Chicago, IL

Transcription Factor (TF) binding sites play a pivotal role in gene regulation networks of plant growth, development, and environmental responsiveness, and thus hold key insights into basic organismal biology. DNA affinity purification sequencing (DAP-seq) offers a highly scalable technique for quickly mapping TF binding sites in a species. We recently introduced multiDAP, a new version of DAP-seq that pools genomic DNA from one or more species in bacteria. Multiplexing reduces technical variability for a given TF across species while increasing the throughput of the assay. Here, we use multiDAP to explore the gene regulatory landscape for the first time in plants, in combination with single-nuclei (sn) ATAC-seq and snRNA-seq data. We propose that the utilization of multiDAP will enable us to identify preserved regulatory elements in plants responsible for crucial biological processes. To do this, we interrogated TF binding sites across 10 plant genomes with 430 TFs, 245 TFs from Arabidopsis thaliana, and sets of 37 TF orthologs from 5 additional plant species, for an additional 185 TFs. We observed inter-species conservation of regulatory elements significantly more than expected, as well as positive correlations of conservation with features like binding signal strength, motif direction, and peak location preference. Low levels of nucleotide diversity and an overrepresentation of ultraconserved sequences in our conserved TF binding sites suggest strong purifying selection acting on these elements. TF target genes showed significant enrichment of functionally relevant categories of elements that are conserved in at least two species, whereas species-specific elements did not tend to show functional enrichments. These results are consistent with our hypothesis that conserved elements identified by multiDAP play key functions in plants. To further explore TF target gene functions, we integrated multiDAP conservation with snRNAseq and snATAC. We observed higher expression and chromatin accessibility from gene targets with conserved regulatory elements, and identified key cell-type specific TFs and their targets. Overall, our results provide profound insights into the evolutionary histories of all significant plant TF families and clades.

De novo annotation of the wheat pangenome reveals complexity and diversity of the hexaploid wheat pan-transcriptome

Rachel Rusholme-Pilcher[1], Benjamen White[1], Thomas Lux[2], Angéla Juhász[6], Gemy Kaithakottil[1], Susan Duncan[1,3], James Simmonds[3], Hannah Rees[1], Jonathan Wright[1], Josh Colmer[1], Sabrina Ward[1], Ryan Joynson[1,4], Benedict Coombes[1], Naomi Irish[1], Suzanne Henderson[1], Tom Barker[1], Helen Chapman[1], Leah Catchpole[1], Karim Gharbi[1], Moeko Okada[5,16,17], Hirokazu Handa[18], Shuhei Nasuda[19], Kentaro K. Shimizu[5,16], Heidrun Gundlach[2], Daniel Lang[2], Guy Naamati[7], Erik J. Legg[8], Arvind K. Bharti[8], Michelle L. Colgrave[6,9], Wilfried Haerty[1], Cristobal Uauy[3], David Swarbreck[1], Philippa Borrill3, Jesse A. Poland[10], Simon Krattinger[10], Nils Stein[11,15], Klaus F.X. Mayer[2,12], Curtis Pozniak[13], 10+ Wheat Genome Project, Manuel Spannagl^[2], Anthony Hall^[1,14]

1 Earlham Institute, Norwich Research Park, Norwich, NR4 7UH, UK.
2 PGSB Plant Genome and Systems Biology, Helmholtz Center Munich, German Research Center for Environmental Health, Neuherberg, Germany.
3 John Innes Centre, Norwich Research Park, Norwich, NR4 7UH, UK.
4 Limagrain Europe, Clermont-Ferrand, Auvergne-Rhône-Alpes, France.
5 Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland.
6 Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, School of Science, Edith Cowan University, Joondalup, WA, 6027, Australia.
7 EMBL-EBI, Wellcome Genome Campus, Hinxton, Cambridgeshire, UK.
8 Syngenta Crop Protection, Research Triangle Park, NC, USA.
9 CSIRO Agriculture and Food, St Lucia, QLD 4067, Australia.
10 Plant Science Program, Biological and Environmental Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Kingdom of Saudi Arabia
11 Center of Integrated Breeding Research (CiBreed), Georg-August-University, Göttingen, Germany
12 School of Life Sciences, Technical University Munich, Freising, Germany.
13 Crop Development Centre, The University of Saskatchewan, Saskatoon, Canada.
14 School of Biological Sciences, University of East Anglia, Norwich, UK
15 Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Seeland, Germany
16 Kihara Institute for Biological Research, Yokohama City University, Yokohama, Japan

(continue to page 71)

(continued from page 70)

Wheat is the most widely cultivated crop in the world with over 215 million hectares grown annually. However, to meet the demands of a growing global population, breeders face the challenge of increasing wheat production by approximately 60% within the next 40 years. The 10+ Wheat Genomes Project recently sequenced and assembled the genomes of 15 wheat cultivars to develop our understanding of genetic diversity and selection within the pan-genome of wheat. We have developed a wheat pan-transcriptome with de novo annotation and differential expression analysis for nine of these wheat cultivars, across multiple different tissues and whole seedlings sampled at dusk/dawn. Analysis of these de novo annotations has facilitated the discovery of genes absent from the Chinese Spring reference, identified genes specific to particular cultivars and defined the core and dispensable genomes. Expression analysis across cultivars and tissues has also revealed conservation in expression between a large core set of homoeologous genes and identified widespread changes in sub-genome homoeolog expression bias between cultivars. Co-expression network analysis uncovered the impact of divergence of sub-genome homoeolog expression and identified tissue-associated cultivar-specific expression profiles. This work provides both a valuable resource for the wider wheat community and reveals diversity in gene content and expression patterns between global wheat cultivars.

Developing accessible bioinformatics capacity for agricultural researchers

Fiona McCarthy[1], Kai Li[2], Ted Kalbfleisch[2], Kyle Parker[3], Shuyu Liu[3], Caylyn Railey4, Andrew Nelson4, Gustavo P. Schettini[5], Fernando H. Biase[5], Yu Shi[6], Chi Zhang[6], Moustafa Shokrof[7], C. Titus Brown[7], Stephanie McKay[8]

1 University of Arizona, School of Animal Comparative & Biomedical Sciences, Tucson, AZ
2 University of Kentucky, Department of Veterinary Science, Lexington, KY
3 Texas A&M, Department of Soil and Crop Sciences, College Station, TX
4 Cornell University, School of Integrative Plant Science, Ithaca, NY
5 Virginia Polytechnic Institute and State University, School of Animal Sciences, Blacksburg, VA
6 University of Nebraska–Lincoln, School of Biological Sciences, Lincoln, NE
7 University of California Davis, School of Veterinary Medicine, Davis, CA
8 University of Missouri, Division of Animal Sciences, Columbia, MO

The combined decrease in genomic sequencing costs and expansion in applications (e.g., transcriptomics, epigenetics, SNP data) have enabled scientists to apply genomics techniques to an increasingly broad range of species and situations. However, while we can rapidly produce more sequence-based data more cheaply than previously, the bottleneck in applying these techniques now lies in the application of bioinformatics to make sense of the data produced. For most researchers, it is now easier to generate genomic data than it is to manage and analyze the resulting data. Several initiatives have been used to develop accessible bioinformatics capacity (e.g., CyVerse, Galaxy, KBase, etc.) but these platforms still require more technical expertise than is commonly available to many scientists and tend to be software agnostic, making open-source software available without recommending or benchmarking specific workflows commonly used for agricultural genomics. We identified common bioinformatic analyses required for in both plant and animal genomics and developed publicly available analysis workflows. These workflows utilize publicly available software and provide extensive documentation along with worked examples. They include workflows for (1) integrating short read and nanopore data; (2) quantifying RNA splice variants; (3) identifying variants and (4) predicting anti-sense lncRNAs. In addition, we have developed and tested a prototype which enables rapid and reference-free analysis of distributed genomics and genetics data sets. Future work includes beta-testing of these workflows with additional data sets and review of the documentation to ensure that these products are adaptable and reproducible. Together, these workflows can provide a core of well documented and accessible analyses to support agricultural researchers, and this process is easily adaptable to additional analyses. Moreover, this core set can provide the bioinformatics community with benchmarked workflows which support the development and comparison of new workflows.

Encouraging reproducibility: A bioinformatic pipeline for scaling single-cell/single-nuclei RNA-seq analyses

Caleb Stull[1,2], Robert Schnabel[1,2]

1 University of Missouri, Genetics Area Program, Columbia, MO
2 University of Missouri, Division of Animal Sciences, Columbia, MO

Single-cell/Single-nuclei (sc/sn) RNA-seq datasets capture a snapshot of cell type compositions and transcriptional profiles. Workflows for sc and sn analyses developed for human and mouse studies are being extended to livestock species. Downstream analyses such as Differentially Expressed Gene Analysis and Bulk RNA-seq deconvolution rely on accurate cell type annotation and increasing sample sizes to maximize statistical power. Publicly available datasets are growing, with heterogeneity in the number of cells and reads per cell targeted. These points illustrate the need to evaluate workflows for livestock species and reproducible cell annotation of integrated datasets. We present a bioinformatic pipeline to reproducibly scale sc/snRNA-seq analysis for individual and integrated datasets. Raw sequence reads are aligned utilizing Cellranger (6.1.2). Quality Control (QC) for ambient RNA, background RNA, and doublet detection is performed with CellBender (0.3.0), scDblFinder (1.17.0), and Seurat (5.0.0). Clustering and marker gene detection are performed with Seurat. Initial cell type annotations are generated with scMayoMap (0.2.0) and SingleR (2.4.1). Sub-clustering and manual annotation can be performed within the framework of the pipeline. The pipeline runs from alignment to annotation for each sample and the integrated dataset. Re-analysis can be achieved easily by updating pipeline parameters. We illustrate how manual decisions on workflow steps can alter cell type annotations. We, therefore, provide guidelines for setting pipeline parameters for QC, normalization, selection of variable features, clustering resolution, and marker gene selection to encourage reproducibility. Finally, we report the cellular landscape of six Bos taurus liver samples with raw sequencing reads generated using 10x Genomics 3' library preparation and sequenced on the Illumina NovaSeq platform. The integrated dataset includes ~18,000 cells and 8 major cell types, including hepatocytes, endothelial cells, cholangiocytes, hepatic stellate cells, B cells, T cells, NK-like cells, and Kupffer cells. This pipeline provides a means for reproducibly scaling sc/snRNA-seq analyses using project-generated and publicly available datasets.

Integrating single-cell genomics into agricultural G2P research: Establishing a FAIR data ecosystem and computational tools for single-cell RNAseq data analysis

Muskan Kapoor[1], Enrique Sapena Ventura[2], Amy Walsh[3], Alexey Sokolov[2], James Koltes[1], Timothy Tickle[4], Wesley Warren[5], Galabina Yordanova[2], Peter Harrison[2], Christine Elsik[3], Tony Burdett[2], Christopher Tuggle[1]

1 Bioinformatics and Computational Biology Program, Department of Animal Science, Iowa State University, Ames, IA
2 EMBL-EBI, Wellcome Genome Campus, Hinxton, Cambridgeshire, United Kingdom
3 Division of Animal Science and Division of Plant Science and Technology, Animal Science Research Center, University of Missouri-Columbia, Columbia, MO
4 Data Sciences Platform, The Broad Institute of MIT and Harvard, Cambridge, MA
5 Division of Animal Science, Life Sciences Center, University of Missouri-Columbia, Columbia, MO

The agriculture genomics community has numerous data submission standards available but little experience describing and storing single-cell (SC, e.g., scRNAseq) data. Other SC-genomics infrastructure efforts, such as the Human Cell Atlas Data Coordination Platform (HCA DCP), have resources that could benefit our community. For example, the HCA DCP is integrated with Terra, a cloud-native workbench for computational biology developed by Broad, Verily, and Microsoft that houses world-leading tools for scGenomics analysis. We described a pilot-scale project that determines whether the current metadata standards for livestock can be used to ingest scRNAseq datasets into Terra in a manner consistent with HCA DCP standards. We further test whether the established resources (e.g., Terra) can be used to analyze the ingested data. The FAANG portal, at EMBL-EBI, has been developed that provides bulk and scRNAseq data access. Importantly, rich scRNAseq metadata can be submitted to FAANG using a semi-automated process, thereby avoiding the need for substantial expert curation to ingest datasets with detailed annotation. We have extended this tool for scRNAseq data so that files can be validated and ingested to using the HCA DCP metadata and data ingestion service and transferred to Terra for further analysis. Once incorporated, datasets will augment the DCP resource for the scientific community. In addition, we also performed Terra-based analysis through Azure and developed a workflow pipeline for the first scRNAseq porcine ingested data. In an extension of these efforts, we tested and developed prototype tools to visualize the output of scRNAseq analyses on genome browsers. This tool allows comparison across tissues and cell populations through the Jbrowse platform. JBrowse now features distinct tracks, showcasing PBMC scRNA-seq alongside two bulk RNA-seq experiments. Notably, the scRNA-seq track shows a bar chart for each gene that illustrates gene expression levels across different cell types. We intend to further build upon these existing tools to construct a scientist-friendly data resource and analytical ecosystem to facilitate SC-level genomic analysis through data ingestion, storage, retrieval, re-use, visualization, and comparative annotation across agricultural species.

Milk metabolites are associated with feed-efficiency traits and are predictive of feed intake in lactating Holstein dairy cattle

Leonora M. James[1], Mary S. Mayes[1], Cori J. Cooper[1], Matthew W. Breitzman[2], Francisco Peñagaricano[3], James E. Koltes[1]

1 Department of Animal Science, Iowa State University, Ames, IA
2 W.M. Keck Metabolomics Research Laboratory, Iowa State University, Ames, IA
3 Department of Animal and Dairy Sciences, University of Wisconsin-Madison, Madison, WI

Feed efficiency traits are pivotal to the dairy industry as feed is the highest cost in production. Genetic tools exist to select for more efficient cattle, but genetic progress is limited due to high phenotyping costs which limit data collection, and thus lower genetic prediction accuracy. Milk data is an underutilized information source that contains metabolites indicative of individual animal metabolism. Millions of dairy cows are milk tested on farm annually. The objective of the study was to: 1) identify metabolites associated with feed efficiency traits; 2) evaluate if milk metabolite data generated on farm was predictive of feed intake. Study 1 utilized 39 cows with extreme residual feed intake (RFI: top and bottom 15%) selected from a larger population of 465 cows with feed efficiency records. Three different assays were utilized to measure metabolites in milk: gas chromatography mass spectrometry (GC-MS), liquid chromatography mass spectrometry (LC-MS), and fatty acids inferred by Fourier Transformed Infrared Spectrometry (FT-IR). A total of 33 GC-MS metabolites ($q<0.20$), 10 LC-MS ontology pathways ($q<0.20$) and 42 fatty acids inferred from FT-IR (Hommel adj-$p<0.05$) were significantly different between RFI groups. Enriched LC-MS metabolite pathways included sphingolipid metabolism, and beta-alanine metabolism pathways ($q<0.20$). Milk FT-IR data from 357 cows were used to predict feed intake in study 2. The best prediction model was the full model including parity, contemporary group, days in milk, and body and milk energy sinks. Alternative models were used together with an across cow validation using a random 25% of cows as the testing set resulting in an R2 of 0.68 and a concordance correlation coefficient of 0.81. Novel metabolites identified in study 1 are candidate indicator traits for feed intake. Milk FT-IR data are routinely collected on US dairy farms, thus the ability to predict feed intake as described in study 2 could allow milk FTIR to be used as an indicator trait or management tool to improve feed efficiency. Both studies require validation but provide promising information that may help select for more efficient and sustainable dairy cows.

Transfer learning and meta-analysis strategies for optimizing genomic prediction accuracy without data sharing

Blessing Olabosoye[1], Gustavo de los Campos[2,3], Jack Dekkers[1], Hao Cheng[4], Juan P. Steibel[1]

1 Iowa State University, Animal Science, Ames, IA
2 Michigan State University, Epidemiology & Biostatistics, East Lansing, MI
3 Michigan State University, Statistics & Probability, East Lansing, MI
4 University of California, Davis, Animal Science, Davis, CA

Genomic prediction (GP) and Genome-wide association (GWA) studies in agricultural species require extensive datasets, often challenging for a single entity to compile. However, the imperative of data sharing among individuals encounters the dilemmas of privacy and intellectual property protection, particularly in assessing hard-to-measure traits such as feed efficiency. To address these challenges, Transfer learning (TL) approaches emerge as promising solutions facilitating collaborative efforts while safeguarding data privacy. We explored TL, including TL-GBLUP and TL-Ridge regression (TL-RR), alongside a meta-analysis approach. The TL approach involves performing GBLUP on each source population, transferring the resulting marker effect estimates to the target population, and then retraining on the target population. This approach is more promising if the source populations (in total) have more individuals than the target population. Implementation was carried out using the gwaR package in R, leveraging publicly available datasets from three pig populations: P1 (N=1920), P2 (N=948), and P3 (N=1237), each with 37069 SNP markers. The evaluation focused primarily on meat quality traits, such as ultimate pH. The accuracy of GP in the target population was investigated by comparing the predicted phenotypes with and without TL to the observed phenotypes in ten-fold cross-validation. Results demonstrated that TL-GBLUP can achieve better accuracy (☒ 4%), but results varied across traits. With TL-RR, the target population marker effects were estimated using a weighted sum of the ridge estimators from the target and source populations. The meta-analysis algorithm required sharing four basic variables: estimated genetic variance, estimated marker effects and their variances, and sample size. The meta-analysis resulted in similar accuracies as TL, but it was particularly useful for genetically stratified populations. This study underscores the potential for significant enhancement of GP by transferring marker effect estimates between similar populations. Consequently, TL and meta-analysis approaches hold promise for improving the predictive accuracy of GP in scenarios where data sharing is challenging.

This research was supported by the intramural research program of the U.S. Department of Agriculture, National Institute of Food and Agriculture, [AG2PI, 2022-70412-38454 and 2023-70412-41054].

Uncovering putative cis-regulatory regions involved in bovine placental development with single-nuclei multiome sequencing

Kimberly M. Davenport[1,2], Mackenzie T. Lowke[1], M. Sofia Ortega[3], Andrew M. Kelleher[4], Wesley C. Warren[1,5], Thomas E. Spencer[1,4]

1 University of Missouri, Division of Animal Sciences, Columbia, MO, USA
2 Washington State University, Department of Animal Sciences, Pullman, WA, USA
3 University of Wisconsin-Madison, Department of Animal and Dairy Sciences, Madison, WI, USA
4 University of Missouri, Department of Obstetrics, Gynecology, and Women's Health, Columbia, MO, USA
5 University of Missouri, Institute for Data Science and Informatics, Columbia, MO, USA

Pregnancy loss in cattle poses a significant challenge for both beef and dairy farms, with only around 30% of cattle successfully carrying pregnancies from fertilization to calving. The intricate processes crucial for successful pregnancy, including proper placental formation, play a pivotal role. While previous single-cell RNA-seq analyses have uncovered dynamic changes in cell populations and gene expression patterns during bovine placental development, the precise contributions of genetic regulatory regions to these changes and their impact on successful pregnancy remain elusive. This study aims to delineate putative regulatory regions across diverse cell populations in the bovine placenta. To achieve this, early (day 40, n=3) and late (day 170, n=2) gestation placental tissues were dissected, and nuclei (> 10,000 per sample) were isolated from cotyledon tissue. Utilizing the 10X Genomics Platform, single-nuclei multiome sequencing (snRNA-seq and snATAC-seq) was conducted. The obtained data underwent analysis with Cell Ranger ARC v2.0.2, Seurat v5.0, Signac v1.12.0, Monocle3, and TFBSTools. Validation of mRNA expression defining cell populations was performed using RNA in-situ hybridization. The results revealed distinct trophoblast, mesenchyme, endothelial, immune, and epithelial cell populations characterized by unique gene expression and chromatin accessibility signatures throughout gestation. ATAC-seq peaks defined open chromatin regions, aiding in the identification of transcription factor binding sites and networks. Notably, transcription factors such as JUN, TEAD4, GATA3, TFAP2A, and TFAP2C, known for their involvement in trophoblast differentiation in other mammalian species, were identified. Putative cis-regulatory regions were defined through co-accessibility and expression networks, revealing variation among cell populations. This study establishes a foundation for understanding gene regulation and expression in the placenta, offering insights into the mechanisms governing pregnancy loss and fertility in cattle. Ultimately, such knowledge holds the potential to enhance fertility, contributing to the profitability and sustainability of beef and dairy operations globally.

Confidence intervals for cross-validation statistics in genomic prediction

Matias Bermann[1], Andres Legarra[2], Alejandra Alvarez Munera[1], Daniela Lourenco[1], Ignacy Misztal[1]

1 Department of Animal and Dairy Science, University of Georgia, 425 River Road, 30602, Athens, GA, USA
2 Council on Dairy Cattle Breeding (CDCB), Bowie, Maryland 20716, USA.

Cross-validation by data truncation is a common practice in genetic evaluations to assess the quality of the prediction of the genetic merit of the individuals. Statistics derived from cross-validation could also be used to predict genetic gain. Two of the most used cross-validation methods in genetic evaluations use a single data partition: the LR method and predictivity or predictive ability, which is the correlation between pre-adjusted phenotypes and estimated breeding values divided by the square root of heritability. Both methods compare predictions with the whole data and a partial dataset obtained by removing the information related to a set of validation individuals. Estimated breeding values with the partial dataset are compared against adjusted phenotypes for the predictivity or estimated breeding values with the whole data in the LR method. Confidence intervals for predictivity and the LR method can be obtained by calculating their variation over time, samples (or folds), or bootstrapping. Analytical or frequentist confidence intervals are unavailable for predictivity and the LR method. However, having them would be beneficial to avoid running several cross-validations and to test the quality of the bootstrap intervals. This study aimed to derive analytical confidence intervals for cross-validation statistics used in genetic and genomic evaluations. We derived Wald confidence intervals for the predictivity and statistics included in the LR method (bias, dispersion, ratio of accuracies, and reliability). The confidence intervals for the bias, dispersion, and reliability depend on the relationship and prediction error variances across the individuals in the validation set. We developed approximations for large datasets to reduce the computational cost of obtaining the confidence intervals. We showed the adequacy of the analytical and approximated analytical confidence intervals and compared them versus bootstrap confidence intervals using a simulated example. The analytical confidence intervals were closer to the simulated ones. Bootstrap confidence intervals tend to be narrower than the simulated ones. The approximated analytical confidence intervals were similar to those obtained by bootstrapping. Estimating the sampling variation of predictivity and the statistics in the LR method without replication or bootstrap is possible for any dataset with the formulas derived in this study.

Precision Genetic Technologies

An Efficient Enzyme-Based Library Workflow for Multiplexed Long Read Whole Genome Sequencing

Maura Costello[1], Zac Zwirko[1], Christianto Putra[1], Ashley Silvia[1], Gavin Rush[1], Joseph Mellor[1]

1 seqWell, Inc., Beverly, MA

The increased output of long-read sequencing instruments such as the PacBio Revio platform has quickly improved capacity to multiplex larger numbers of samples in a single sequencing run. This is particularly true for applications that have relatively lower data output requirements, such as microbial whole-genome sequencing or low-pass population genotyping. For these types of applications, the capabilities of instrument output have outpaced the development of efficient library prep workflows that allow users to multiplex large numbers of samples. One of the root causes of this bottleneck is that most long read library workflows still rely on mechanical or hydrodynamic shearing to fragment DNA molecules to the ideal multi-kilobase size that is optimal for long read sequencer loading.

Here, we demonstrate the implementation of a novel tranposase-based technology (LongPlex™) for long fragment multiplexing that is optimized for large scale sequencing workflows for 10s to 100s of samples. The LongPlex product uses a plate-based format of 96 indexed transposome reagents to pre-index and fragment high molecular weight genomic DNA to multi-kilobase fragments that can then be pooled and further adapted with standard long read sequencer-specific adapter chemistry. Among other advantages versus mechanical library approaches, this enzymatic workflow is highly amenable to automation and reduces the number of final libraries that need to be prepared for sequencing.

We show the application of this technology to multiplexed whole genome sequencing in microbial gDNA samples on the PacBio Revio platform, in addition to proof-of-concept data on low-pass sequencing of larger eukaryote gDNA specimens. We further demonstrate the relative performance of PCR-free and limited PCR-based fragment enrichment for a range of gDNA inputs mass levels, and how this corresponds to sequencing coverage uniformity and assembly characteristics.

Comparison of 3D chromosome conformation between parental haplotypes in four cattle tissues

Morgan R. Stegemiller[1], Benjamin D. Rosen[2], Kristen L. Kuhn[3], Holly L. Neibergs[4], Temitayo A. Olagunju[1], Stephanie D. McKay[5], Timothy P.L. Smith[3], Brenda M. Murdoch[1]

1 University of Idaho, Department of Animal, Veterinary, and Food Sciences, Moscow, ID
2 United States Department of Agriculture (USDA), Agricultural Research Service (ARS), Animal Genomics and Improvement Laboratory, Beltsville, MD
3 United States Department of Agriculture (USDA), Agricultural Research Service (ARS), U.S. Meat Animal Research Center, Clay Center, NE
4 Washington State University, Department of Animal Sciences, Pullman, WA
5 University of Missouri, Division of Animal Sciences, Columbia, MO

The three-dimensional (3D) genome organization within the nucleus is vital for many biological processes. The hierarchical 3D organization of the genome is comprised of chromosome territories, compartments ("open" active and "closed" inactive regions), topologically associated domains (TADs), and loops. Chromatin conformation is highly dynamic during mammalian development and is tissue specific. There has been little 3D genome research completed in cattle and even less addressing haplotype-specific organization. A Wagyu x Charolais F1 fetal cross was harvested at 120 days gestation with spleen, muscle, lung, and liver tissue subjected to Micro-C analysis, a variation of chromatin conformation capture (Hi-C) which reveals the position of nucleosomes on chromatin. The Micro-C reads were mapped independently to each parental haplotype of "complete" telomere-to-telomere genome assemblies to establish parental haplotype-specific conformations. Compartments were called using 2.5Mbp per bin and TADs were identified at 50kb resolution. Compartments were analyzed for regions where one haplotype switched compartments and the other did not, revealing 13, 20, 11, and 31 bins compartment switches respectively in spleen, muscle, lung, and liver. Annotation from the reference genome ARS-UCD_1.2 was lifted to both genome assemblies to identify gene ontology (GO) terms from the regions of compartment switches with respect to the haplotype in the active compartment. The top GO term in open Wagyu spleen was lysosome activity (P=8.1E-12) where Charolais spleen is associated with serine-type endopeptidase (P=1.2E-6). Connexin complex is significantly associated with Wagyu lung (P=1.6E-4) and connexins are important in embryonic development and conducted response in microvasculature. A difference in TAD size between the two haplotypes was observed with Wagyu averaging 21kb, 20.7kb and 4.3kb larger in the spleen, muscle, and lung tissue and Charolais TADs being an average 14.8kb larger in the liver. Defining compartment determination and 3D conformation during fetal development can show parts of the genome that are interacting, and which segments of DNA are active or inactive. Assessing conservation of topological regions between haplotypes can indicate critical regions potentially impacting phenotypic variation between breeds.

DEVELOPMENT OF BACTERIAL WILT RESISTANT POTATO USING *Efr, Jim2 and Roq1* GENES

Mary N. Mwangi[1,2,3], Agnieszka Witek[2], Eric Magembe[1], Steven Runo[3], Alex Schultink[4], Jonathan D G Jones[2], Marc Ghislain[1]

1 International Potato Center, Nairobi Kenya.
2 The Sainsbury Laboratory, Norwich Research Park, Norwich, UK
3 Kenyatta University, Dept of Biochemistry, Microbiology and Biotechnology, Nairobi Kenya
4 Fortiphyte Inc., Berkeley, CA, United States.

mary.mwangi@cgiar.org

Potato (Solanum tuberosum L.) in Africa is mainly grown by smallholder farmers. In Kenya, it is the 2nd most important crop, estimated to employ more than 2.5 million people. Bacterial wilt (BW) caused by Ralstonia solanacearum is a devastating disease for 74% of farms in Kenya. The current management strategies have limited efficacy and conventional breeding has had limited success to develop BW tolerant varieties due to lack of sources of stable and broad-spectrum resistance. We hypothesize host resistance using pattern recognition receptors (PRR) and Nucleotide-binding Leucine rich repeat (NLR) proteins has been shown to confer resistance to BW by inducing pathogen triggered immunity (PTI) and effector triggered immunity (ETI) respectively. EFR (Elongation Factor Tu receptor) from Arabidopsis thaliana has been shown to enhance resistance against bacterial wilt. The Roq1 TIR-NLR gene and Jim2, a receptor like cytoplasmic kinase family XII (which signals via Zar1) from Nicotiana benthamiana have been shown to mediate recognition of RipB and RipP1 Ralstonia effectors respectively. In this study we bioengineered the Kenyan promising potato variety 'Unica' using a stack of Efr, Roq1 and Jim2 genes through Agrobacterium-mediated genetic transformation of internodal potato explants. Regenerants were selected on chlorsulfuron to produce putative transgenic events. Molecular characterization selected events positive for Efr, Roq1 and Jim2 genes, negative for left and right border backbone vector sequences and that carries a single T-DNA copy. ROS assays confirmed Efr gene activity. Transiently expression of XopQ and XopJ in transgenic Unica lines confirmed lines carrying functional Roq1 and Jim2 genes. We are now conducting greenhouse bioassays using a local strain of R. solanacearum on these events to evaluate their resistance to bacterial wilt in comparison to the moderately resistant variety Cruza-148 and non-transgenic equivalents.

Key Words
Bacterial Wilt, Genetic engineering, Plant Recognition Receptors, Ralstonia solanacearum

Genome editing for crop improvement in the Global South

Anindya Bandyopadhyay[1], Zhengyu Wen[2], Mark Jung[3], Mohan Achary[1], Kanwarpal S. Dhugga[1], Todd Jones[3], Mario Pacheco[1], James Karanja[4], Suresh Mahabaleswara[1], Yoseph Beyene[1], Ann Murithi[5],Fengzhong Lu[1],Ping Che[3], Huirong Gao[3], Ning Wang[3], Kevin Pixley[1], Sarah Hearne[1]

1 International Maize and Wheat Improvement Center (CIMMYT), Department of Biotechnology, Texcoco, Mexico
2 KeyGene, Rockville, MD
3 Corteva Agriscience, Johnston, IA
4 Kenya Agricultural and Livestock Research Organization, Nairobi, Kenya
5 Iowa State University, Department of Agronomy, Ames, IA

Genome editing offers new possibilities for transforming agricultural systems through development and/or transfer of traits with increased precision and more competitive timeframes compared with traditional breeding. Since its discovery a decade ago, the CRISPR-Cas system has been widely applied in research enhancing the speed of breeding, generating mutations, multiplexing gene targets, SNP changes, and epigenome editing. At CIMMYT, we are working on product development focused on traits of importance for the Global South by directly editing popular elite cultivars of maize and wheat and targeting trait development in opportunity crops such as groundnut, millets, and pigeon pea.

We present a number of opportunities around gene editing in our current portfolio, highlighting public-private partnerships in the application of genome editing. We discuss projects in lowering aflatoxin contamination and improving human health in groundnut, improving shelf-life and opening new market opportunities in pearl millet, and fine mapping and editing for resistance against the scourge of maize lethal necrosis.

Large animal models to advance our understanding of genetic-driven pulmonary vascular disease

Sanjeev A. Datar[1], Austin R. Brown[2], Devon S. Fitzpatrick[2], Oluwafemi A. Falade[2], Josie F.Trott[2], Rachel Hutchings[1], Elena K. Amin[1], Eric G. Johnson[3], Jessica M. Morgan[3], Hythem Nawaytou[1], Omar A. Gonzales Viera[3], Tara Urbano[3], Bret McNabb[3], Eric D. Austin[4], Jeffery R. Fineman[1], Alison L. Van Eenennaam[2]

1 Department of Pediatrics, University of California, San Francisco, CA
2 Department of Animal Science, University of California, Davis, CA
3 School of Veterinary Medicine, University of California, Davis, CA
4 Vanderbilt University School of Medicine, Nashville, TN

Pulmonary Arterial Hypertension (PAH) is a rare vascular disorder of humans occurring in 1-2 people per million individuals. Monoallelic mutations in the bone morphogenetic protein receptor type 2 (BMPR2) gene are the most common genetic risk factor, with about 20% of carriers affected. Biallelic BMPR2(-/-) mutations are embryonic lethal. BMPR2 is a serine/threonine kinase receptor in the transforming growth factor beta superfamily. Although autosomal dominant, BMPR2(+/-) has variable penetrance, pathologically expressing in ~42% of females and 14% of males. Despite well-known associations with PAH, BMPR2`s role in the disorder's progression remains poorly understood due in part to the lack of appropriate preclinical genetic models. The goal of this project was to develop a high-fidelity large animal pre-clinical model of PAH by creating heterozygous BMPR2(+/-) knockout sheep using a CRISPR/Cas9 gene-editing system in sheep embryos to target exon 3 of ovine BMPR2. To overcome the embryonic lethality of BMPR2(-/-), a single-stranded oligodeoxynucleotide (ssODN) that disrupted the PAM site was designed to serve as a synonymous repair template. Ovine zygotes were electroporated six hours post-fertilization using sgRNA (100 ng/µL)/Cas9 (200 ng/µL) ribonucleoprotein complexes and ssODN (600 ng/ul). Five or six edited blastocysts were transferred to each of 8 recipient ewes resulting in 4 liveborn lambs (3♀:1♂) in August 2023. All of the lambs were BMPR2 heterozygotes with indels resulting in nonsense mutations in one BMPR2 allele, and a wildtype allele that carried the silent disrupted PAM sequence of the ssODN sequence providing evidence of template-guided repair. Females phenocopied PAH disease progression. Semen was collected from the ram, and segregation of both the wildtype and 49 bp deletion alleles was confirmed. This ram is now being crossed with wildtype ewes to generate a steady supply of BMPR2(+/-) heterozygous and half-sibling wildtype BMPR2(+/+) control lambs. This genome editing approach enabled the development of a dependable source of BMPR2(+/-) lambs that the human medical community can use to interrogate the natural history of PAH disease progression in a manner that is not feasible in human patients or rodent animal models, including what conditions

Leveraging nanopore adaptive sampling to inform cattle precision breeding

Emily Kiugu[1], Nicole Flack[1], Evan Kipp[1], Luciano Caixeta[2], Brian Crooker[3], Tad Sonstergard[4], Peter Larsen[1]

1 University of Minnesota, Department of Veterinary Biomedical Sciences, St. Paul, MN
2 University of Minnesota, Department of Veterinary Population Medicine, St. Paul, MN
3 University of Minnesota, Department of Animal Science, St. Paul, MN
4 Acceligen™, Eagan, MN

Precision breeding in cattle involves making precise DNA modifications to achieve a phenotype having specific traits of interest in the next generation. The foundation for gene editing is the ability to identify and confirm target genes and gene variants using sequence data. Recent bioinformatic pipelines offered by Oxford Nanopore Technologies (ONT) allow for the selective sequencing of specific genomic targets across a given genome. Remarkably, this approach is purely bioinformatic and is conducted in real-time as sequencing is occurring, with DNA molecule "accept" or "reject" decisions made at the level of individual sequencing pores. Here we demonstrate how the ONT adaptive sampling technique can be leveraged to simultaneously target hundreds of loci across the Bos taurus genome, including specific regions critical to precision breeding efforts in cattle. We targeted 237 genes associated with thermoregulation and immune response in cattle using the latest Bos taurus genome assembly. Genomic DNA was extracted from a Holstein dairy cow followed by library preparation and sequencing using ONT's MinION sequencer. Our results indicate that adaptive sampling was successful as it enriched 98% of our 237 targeted regions at a median coverage depth of 11.5x (3.4-fold above background). Additionally, we present data on phased nucleotide variants and differential DNA methylation facilitated by ONT long reads and direct DNA modification detection, respectively. We posit that the ONT adaptive sampling approach will serve as a valuable tool for the rapid and precise assessment of multiple loci of interest targeted for gene editing in cattle.

Precision genome editing for crop improvement

Feng Zhang[1,2]

1 Department of Plant and Microbial Biology, University of Minnesota, Minneapolis, MN
2 Center for Precision Plant Genomics, University of Minnesota, Minneapolis, MN

Recent advances in sequence-specific nuclease technologies, especially the CRISPR/Cas system, have made it possible to introduce targeted genetic modifications across a wide range of plant species. In recent years, while substantial efforts have been made to improve precise genome editing, achieving high efficiency remains challenging in most plant species. To address this challenge, my lab investigated the intricate interplay among CRISPR gene editing, DNA repair mechanisms, and epigenetic factors. Our research has revealed that CRISPR-Cas9, initially believed to primarily induce blunt cleavages, frequently generates staggered cleavages in plants. Moreover, we demonstrated that these CRISPR-induced staggered cleavages can be effectively utilized to facilitate the generation of precise small and large DNA insertions. Furthermore, through the screening of DNA repair mutants in Arabidopsis, we have identified the pivotal gene responsible for precise DNA end processing. Modulation of its expression could substantially enhance targeted insertion frequency. Building on these discoveries, we developed a new targeted insertion approach that enabled us to precisely engineer cis-regulatory elements for disease resistance in rice. The edited rice lines exhibited strong resistance to infection by the pathogen Xanthomonas oryzae pv. oryzae in an inducible and strain-specific manner. Taken together, our research holds the potential to pave the way for achieving efficient gene editing with improved precision for crop improvement.

Topological associated domains identification in the cattle genome

Loan T. Nguyen[1], Jung Hyungtaek[2], Jun Ma[3], Stacey Andersen[3], Elizabeth Ross[1]

1 Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Brisbane, Queensland, Australia
2 Australian National University (ANU) College of Health and Medicine, Australian National University, Canberra, Australia
3 Genome Innovation Hub, The University of Queensland, Brisbane, Queensland, Australia

Recent molecular technology advancements have significantly enhanced our ability to scrutinize the cattle genome with unprecedented accuracy. While numerous studies have delved into genetic variations and the transcriptional landscape of the bovine genome, the exploration of its topological associating domains (TADs) has been limited, primarily due to the high costs of necessary assays. Before this research, the only TAD data for cattle were based on synteny maps from humans, mice, and dogs. This study used Pore-C technology for genomic analysis and identifying TADs in cattle, focusing specifically on the Brahman breed. Pore-C streamlines the library preparation process compared to traditional methods like Hi-C, by eliminating steps such as biotin labelling and fragmentation. This technique is particularly efficient in capturing multi-way chromatin interactions, with over 50% of reads identified in this study showing contact orders of three or higher. Additionally, this project also provided experimental TAD structures in cattle, using a high-quality Brahman cattle genome assembly (UQ_Brahman_2). A total of 3,036 and 3,094 TADs across 30 chromosomes were identified for input Brahman and ARS-UCD1.2 assemblies, respectively. When compared to TADs reported in previous studies, the results reveal a low level of concordance, pointing to variations that could result from species and breed differences, as well as differences in computational analysis methods. The study's insights are valuable for understanding the regulatory systems and genomic variations across species, contributing to our knowledge of bovine genetics and its implications in breeding and the study of complex traits.

Unlocking Innovations From Basic Science

Andropogoneae crop wild relative (CWR) climatic niche diversity and environmental gap identification of heat stress tolerance using species distribution modelling

WiL Arthur Harris[1], Ian J. Wright[2], Rachael V. Gallagher[2], Mark Cooper[1]

1 University of Queensland, Queensland Alliance for Agriculture and Food Innovation, Centre of Excellence in Plant Success in Nature and Agriculture, St. Lucia, QLD, Australia

2 Western Sydney University, Hawkesbury Institute for the Environment, Centre of Excellence in Plant Success in Nature and Agriculture, Richmond, NSW, Australia

Global agricultural sectors are increasingly threatened by climate change. Large areas of Australia, especially in the fertile east coast grain basket, are predicted to experience elevated temperature and precipitation stress. Given that, currently, most calories in human diets globally derive from members of the grass family Poaceae. Therefore, expanding genetic and phenotypic diversity in grain crops to help mitigate the negative effects of climate change on yields is imperative. The Andropogoneae grasses of the PACMAD clade are of specific interest for crop improvement. The tribe Andropogoneae includes the 3rd and 5th most important food crops (Zea mays and Sorghum bicolor), the 2nd most valuable luxury crop, and the foremost biofuel crops (Saccharum officinarum and Miscanthus sinensis). Their use of NADP-ME C4 photosynthesis reduces photorespiratory carbon losses at high temperatures, which gives advantages in adaptability to hotter and drier environments compared to C3 crops like rice. Despite these advantages as C4 plants, the complexity of Genotype by Environment by Management (GxExM) interactions in agricultural crop production, along with variable levels of genetic erosion within the crop species of Andropogoneae, motivates seeking novel crop improvement opportunities. In addition, temperatures in the growing regions may exceed critical limits for reproductive capacity before photosynthetic limits are hit. Bioprospecting in crop wild relatives (CWRs) is a potential source of novel crop improvement. This improvement can take forms such as allele identification and introgression through genetic modification or breeding, known also as allele mining. The niches of CWRs were modelled within six closely related genera of the Saccharinae (containing Sugarcane and Sorghum) and Tripsicanae (containing Maize) subtribes to identify potential tolerance breadths to heat-related climatic variables. Combining species distribution models, SDMs, and predicted niche occupancy profiles, candidate taxa that may be suitable for introgression and targeted allele mining identified as sources for novel heat stress tolerance resistance were identified. The information gathered through PNO comparison will inform future phenotyping by identifying promising taxa.

High-throughpout Gene Editing to Identify Mechanisms of Mycorrhizal Symbiosis Regulation

Evan E. Ellison[1], Giles E.D. Oldroyd[1]

1 Crop Science Centre, University of Cambridge, Cambridge, United Kingdom

Developing crop varieties that take full advantage of native environmental resources is critical for sustainable food production. One approach to increase plant resilience and reduce fertilizer input is by improving plant associations with arbuscular mycorrhizal fungi (AMF). AMF are ubiquitous in global agricultural soils and enhance nutrient and water capture in exchange for carbon from the host. This symbiosis, however, is restricted by the application of excess inorganic nitrogen and phosphorus. Genomic and transcriptomic resources have provided large data sets of candidate genes underlying nutrient mediated regulation and manipulation of these networks can override symbiosis suppression. We are interested in characterizing modifications to nutrient regulatory networks and their impact on AMF engagement and plant productivity.
To facilitate crop improvements needed in a rapidly changing climate, reverse genetic technologies must be integrated into breeding-like programs of selection and line advancement. We are employing viral vector mediated gene editing to generate mutants in suites of genes related to AMF symbiosis and nutrient homeostasis. Using rapid targeted mutagenesis and machine learning assisted high-throughput phenotyping, we are able to quickly screen candidates for the ability to regulate symbiosis. This system enables continuing line improvement by mutant reinfection; stacking combinations of alleles to create diverse populations that have differential engagement with AMF under symbiosis limiting conditions. We anticipate our iterative gene editing and line selection approach will allow us to identify combinations of alleles that can be utilized in major crops species to fine-tune the relationship between inorganic fertilizer uptake and symbiotic benefit.

Mapping wheat reproductive development with spatial transcriptomics

Katie Long[1], Ashleigh Lister[2], Chen Ji[1], Yuanrong Pei[1], Anna Backhaus[1], Nicola Trozzi[1], Xumei Luo[3], Jun Xiao[3], Iain McCaulay[2], Wilfried Haerty[2], Richard Smith[1], Philippa Borrill[1], Cristobal Uauy[1]

1 John Innes Centre, Norwich Research Park, Norwich, United Kingdom
2 Earlham Institute, Norwich Research Park, Norwich, United Kingdom
3 CAS-JIC Centre of Excellence for Plant and Microbial Science, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China

At the beginning of its development, the wheat spike (the grain-bearing inflorescence) starts as relatively uniform meristematic tissue. From this, complex structures arise including the floral organs, guided in part by spatially restricted gene expression and differential rates of cellular growth. Through emerging technologies in spatial transcriptomics such as MERFISH (multiplexed error-robust fluorescence in situ hybridisation), we can map the transcriptome to cellular resolution in young wheat spikes to understand the genetic regulation of cellular identity. To complement this, live imaging and growth analysis with MorphoGraphX in young wheat spikes characterises the differential rates of cellular growth during development, further establishing the mechanisms shaping the final form of the spike.

Single-cell transcriptome analysis of channel catfish whole spleen and IgM-positive splenic fractions provides insight into the fish immunome from an aquaculture-relevant species

Johanna Aldersey[1,2], Miles D. Lange[2], Benjamin H. Beck[2], Jason W. Abernathy[2]

1 Oak Ridge Institute for Science and Education (ORISE), ARS Research Participation Program, Oak Ridge, TN
2 United States Department of Agriculture, Agricultural Research Service (USDA-ARS), Aquatic Animal Health Research Unit, Auburn, AL

The catfish industry is the largest sector of U.S. aquaculture production. Additionally, channel catfish, Ictalurus punctatus, is an important model for studying teleost fish immunity. Given its role in food production, the catfish immune response to industry-relevant pathogens has been extensively studied and has provided crucial information on innate and adaptive immune function during disease progression. To further examine the channel catfish immune system, we performed single-cell RNA sequencing on both whole spleens (n=3) and IgM-positive splenic B-cells (n=3). Spleen cell suspensions were prepared by passing tissues through a cell sieve. For IgM-positive B-cells, spleen lymphocytes were isolated and then labeled with a recombinant monoclonal antibody that is reactive to catfish IgM and sorted via flow cytometry. Single-cell RNAseq libraries were then prepared using the 10X Genomics Chromium X with the Next GEM Single Cell 3' Reagents and sequenced on an Illumina NovaSeq X Plus. Each demultiplexed sample was aligned to the CoCo_2.0 channel catfish reference assembly, filtered, and counted to generate feature-barcode matrices. From whole spleen samples, outputs were analyzed both individually (SP1, SP2, and SP3) and as a normalized aggregate. SP1, SP2, and SP3 generated 198,304,083-422,617,575 reads from an estimated 5,140-11,463 cells with approximately 25,594-41,874 reads/cell. The reads aligned to 22,234-23,144 genes and 59–80% of reads had a valid barcode and mapped confidently to the transcriptome. The median number of genes found per cell was 554-733. The aggregated cell data had 688,285,722 reads from an estimated 24,351 number of cells and approximately 28,265 reads/cell. 73% of reads had a valid barcode and mapped confidently to the transcriptome and a median of 664 genes were found per cell. K-means analysis on individual samples found nine groups with significantly altered gene expression. Initially, two major clusters were identified including erythrocytes (2,516–4,934 cells) and B-cells (405–2235 cells). Additional gene expression profiles within whole spleen samples as well as IgM-sorted data analytics will be presented and discussed. These data will serve as the foundation for a spleen cell atlas of this primary immune tissue in fish and allow for profiling of B-cell subsets to a high-resolution.

From the Cutting Edge - New Ideas and Opportunities

Biological nitrification inhibition in corn stover: How does variation in stover biochemical composition affect nitrogen loss in soil?

Sam Herr[1], Ela Jusino Montalvo[2,3], Bryan Emmett[2], Amparo Sofía Daza Castañón[4], Natalia de Leon[4], Sakiko Okumoto[5], Elizabeth A. Ainsworth[6], M. Cinta Romay[7], Edward S. Buckler[1,7,8], Michael A. Gore[1]

1 School of Integrative Plant Science, Plant Breeding and Genetics Section, Cornell University, Ithaca, NY
2 United States Department of Agriculture, Agricultural Research Service (USDA-ARS) National Laboratory for Agriculture and the Environment, Ames, IA
3 Oak Ridge Institute for Science and Education, US Department of Energy, Oak Ridge, TN
4 Department of Plant and Agroecosystem Sciences, University of Wisconsin-Madison, Madison, WI
5 Department of Soil and Crop Sciences, College of Agriculture and Life Sciences, Texas A&M University, College Station, TX
6 Department of Plant Biology, University of Illinois Urbana Champaign, Urbana, IL
7 USA Institute for Genomic Diversity, Cornell University, Ithaca, NY
8 United States Department of Agriculture, Agricultural Research Service (USDA-ARS), Ithaca, NY

Biological nitrification inhibition (BNI), a process in which plant metabolites inhibit the microbial conversion of ammonium to nitrate in soil, is a promising plant trait to reduce nitrogen loss in agricultural fields. Recently, metabolites with BNI activity were identified in the root exudates of corn (Zea mays L.). Yet because a substantial portion of the nitrogen is lost in early spring before the root system is established, the potential impact of corn root exudates to reduce nitrogen loss could be low. Interestingly, metabolites with BNI activity, such as vanillin, that have been identified in other species, are closely linked to secondary cell wall formation and have been found within different tissue types. Herein, we propose to explore how the biochemical composition of corn stover, the aboveground biomass remaining after grain harvest, affects the nitrification pathway during decomposition in the soil. In the first year of our study, we set out to explore the potential of BNI compounds within corn tissue. We sampled stalk and leaf tissue at maturity from a diverse collection including hybrids from the Genomes to Fields (G2F) initiative and inbred founders of the maize nested association mapping (NAM) panel. We developed a high throughput in vitro BNI assay to test the inhibitory effect of water-soluble compounds extracted from ground tissue on nitrifying bacteria. We will then quantify metabolites from extracts with high-performance liquid chromatography to identify associations with BNI activity. To associate the water-soluble compounds with the cell-bound metabolites, we performed pyrolysis molecular beam mass spectrometry to quantify the relative lignin composition of the stalk tissue. This combined approach is expected to dissect the relationship between stover composition and metabolites with BNI activity that could be leached in the initial stages of tissue decomposition. We hypothesize that lignin abundance will correlate with the BNI activity of the aqueous extract. Future research will use corn genotypes based on the results of the combined approach to test the effect of stover decomposition on nitrous oxide emissions in agricultural fields. Overall, our research can help determine the utility of breeding for stover composition to reduce nitrogen loss in agricultural fields.

Human microbiome profiling of nasal, throat, and skin swabs demonstrating significant reduction in chimeric 16SrRNA sequence formation for improved accuracy of metagenomic

Chris Streck[1], Yann Jouvenot[1], Akshit Kanda[1], Joseph Jin[1], Allan Yu[1,] Fareed Nabkel[1], Yassine Kabouzi[1], Pranav Patel

1 n6 Tec, Inc., Pleasanton, CA

Bacterial population diversity in human microbiome samples is now commonly used to characterize the interaction between host and localized flora across a range of normal and diseased states. Technical challenges in the preparation of both PCR-based and shotgun metagenomic sequencing have been well documented to reduce accurate representation of populations, including chimeric products created through hybrid parent sequences resulting in false identification of novel species and an overall inaccuracy of the true diversity. The production of these chimeras have been documented to be reproducibly generated across a number of studies generating false identification of low-abundant sequences at over 70% rates. Bioinformatics tools have been developed to detect some of these events, but are still hampered by the presence of the chimeras generated in the initial stages of the experiment. Here we show an improved method for library generation using individually controlled PCR to reduce or eliminate the initial creation of the hybrid events.

In this study, human swabs were collected from nasal, throat and skin samples and prepared using a universal PCR primer set targeting the 16S rRNA gene after the V9 region. Amplicons were generated targeting the V3 and V4 region with an improved amplification method using individually controlled reactions for each well (iconPCR). Samples were subsequently sequenced on both short-read (Illumina, Element) and long-read (PacBio) sequencing platforms. The comparative results show a significant reduction in the overall chimeric transcripts and in improved representation of the true microbiological diversity.

Innovating Genotyping Using Long-Read Low-Coverage Sequencing

Kendall Lee[1], Josh Clevenger[1], Walid Korani[1]

1 HudsonAlpha Institute for Biotechnology, Huntsville, AL

Long-read low-pass (LRLP) sequencing-based genotyping is the future of highly effective genotyping. Short reads are on average a length of ~150 base pairs, whereas Pacific Bioscience's (PacBio) long-read sequencing technology allows us to sequence DNA as long as 20,000 base pairs. LRLP genotyping offers an abundance of benefits for plant and animal breeding that short-read sequencing cannot achieve including improved mapping resolution and identification of important large structural variants (SVs). Here we present work being carried out to achieve affordable and high-throughput LRLP genotyping, including the optimization of high-throughput wet lab protocols and analysis of early LRLP data in peanuts. High-throughput extraction of high-molecular-weight DNA has been the largest hurdle to achieving affordable LRLP sequencing. PacBio has recently successfully tested a high-throughput extraction protocol in humans and has partnered with us to optimize this protocol for plants. We have succeeded in performing plate-based extractions and are optimizing downstream processing, including shearing and library prep. We have early results for peanut samples that were sequenced on PacBio's Revio sequencing platform as a pool. The genotyping results of these samples were compared to the short-read-based genotyping of the same samples. Results show improved coverage and the ability to call large SVs. In addition, our confidence in alignment increases to over 90% as the read length reaches 12.8kb, while the number of filtered SNPs available continues to increase as the read length increases. In contrast, short-read methods result in almost 50% SNP data loss due to a lack of confidence in alignment. Retention of vital SNPs and the ability to call large SVs make LRLP genotyping able to pinpoint traits of interest that were previously inaccessible to us.

Parentage validation and pedigree reconstruction of South African beef cattle using genomic information

Yandisiwe P. Sanarana[1,2], Azwihangwisi Maiwashe[2] , Donagh P. Berry[1,3], Cuthbert Banga[2], Este van Marle-Köster[1]

1 University of Pretoria, Department of Animal and Wildlife Science, Pretoria, Hatfield, 0002, South Africa
2 Agricultural Research Council-Animal Production, Pretoria, Irene, 0062, South Africa
3 Teagasc, Animal and Grassland Research, and Innovation Center, Fermoy, Cork, P61 C996, Ireland

Genomic data on individual animals is now becoming routinely available in South Africa. The objective of the present study was to validate recorded ancestry of Bonsmara (BON) and Drakensberger (DRB) animals but also discover unknown genomic relationships using the single nucleotide polymorphism (SNP) markers. Genomic relationships were validated based on the count of autosomal Mendelian conflicts between the parent-progeny pairs, which were sequentially used to detect possible regions with hemizygous deletion (hDEL) in the genome. A high frequency of Mendelian conflicts was observed in chromosomes 1 and 6 in the BON and DRB, respectively. On average, 8.5% and 10.1% of parentage errors were observed in the pedigrees of the BON and DRB, respectively. Errors detected in dam-progeny pairs varied from 7.0% (DRB) to 7.6% (BON), whereas for the sire-progeny pairs, error rates were relatively higher and ranged from 10.0% (BON) to 17.2% (DRB). SNP-based parent-progeny discovery was possible for 69 relationships which were not previously recorded in the parentage information. This included 19 and 20 sire-offspring pairs; 12 and 18 dam-offspring pairs as well as 4 and 9 half-siblings of the BON and DRB. The suspected regions to contain deletions based on the approach of using the SNP-chip data and Mendelian mismatches between the parent-progeny pairs were detected from 11 regions of 10 chromosomes (3, 6, 7, 9, 10, 11, 17, 24, 25 and 28). These suggestive regions were traced from the most used breeding animals, which were 7 (BON) and 5 (DRB) bulls, 2 (BON) and 7 (DRB) cows. This study demonstrated how SNP markers can be utilized to confirm and determine parentage; they can also be used to identify locations in carriers' genomes that are linked to hDEL. It is advised to use such methods to genotype database accumulations to enhance genetic and genomic analyses. Although the SNP density of 91 185 SNPs used in the present study provided a high precision in performing these analyses, a much lower density panel is desirable in practice to perform parentages. On the other hand, a large familial whole genome SNP-chip based data is necessary to identify the regions of homozygosity. Due to the differences of the required properties SNPs, these analyses may be undertaken independently.

Keywords: Genomic evaluations, genealogy, hemizygosity, pedigree discovery, half-siblings

Seagrass genomics to inform the design of innovative saline crops

1 Pramod Pantha, Louisiana State University

Seagrasses are marine flowering plants that have independently evolved to survive high salinity. We sought to examine the genomic traits that facilitated the transition from freshwater to marine habitats using the tropical seagrass Halophila stipulacea and its closely related freshwater species Vallisneria americana in Hydrocharitaceae. We generated the draft genomes of H. stipulaceae and V. americana and conducted a comparative genomic analysis with thirty-four additional diverse angiosperms to investigate how convergent genomic features may have facilitated the selection of desirable traits in colonization of marine environments. We examined eleven fully submerged, partially submerged, and amphibious genomes in addition to crops and commonly used plant models to capture recurrent gene loss or expansion that are shared among salt-tolerant species. The H. stipulaceae genome had an assembled size of 6 GB with 35,523 protein-coding genes, while the V. americana genome was 11 Gb and contained 81,237 protein-coding genes. Despite a large difference in their genome sizes among seagrasses, our results indicate significant shared gene losses associated with developmental processes and stress tolerance between H. stipulaceae and Zostera marina. Remarkably, the two seagrass species showed gene loss in the same pathways, but for different biosynthesis or regulatory steps. Many of the ABA and auxin regulatory genes associated with salt stress had uniquely shared features in the seagrasses. Gene families associated with photosynthesis, sugar transport, and reproductive development had increased copy number in seagrasses compared to their terrestrial sister species. The seagrass genomes also showed similarities in gene family diversification related to UV and osmotic stresses shared with select marine algal genomes likely highlighting an evolutionary reversal back to the sea. Collectively, these genomic features highlight some of the recurrent evolutionary successful genetic adaptations that can be explored when designing novel saline adapted crops to fit current agricultural needs. This becomes an imperative necessity in seeking solutions to boost crop production amidst a climate crisis, especially when our current elite crops encounter complete failure or substantial yield losses under high salinity conditions.

Using non-invasive microbiome samples to select low-methane-emission cattle

Chian Teng Ong[1], Ziming Chen[1], Elizabeth M. Ross[1]

1 The University of Queensland, Queensland Alliance of Agriculture and Food Innovation, Centre of Animal Science, Brisbane, Queensland, Australia

As an integral aspect of their digestive process, ruminants, including cattle, generate greenhouse gases (GHGs), notably methane. In the year 2022, livestock farming contributed to a substantial 11.1–19.6% of total GHG emissions, solidifying its position as one of the significant anthropogenic sources of GHGs. Enteric methane, a prevalent by-product of livestock farming, significantly contributes to global warming. The escalating global population, coupled with an increasing demand for protein-based diets, anticipates the continued expansion of livestock ranching, intensifying concerns about the adverse environmental impact caused by methane emissions. To counteract these detrimental effects, concerted research efforts have been directed towards fostering more sustainable farming practices, dietary adjustments for livestock and the exploration of innovative technologies designed to capture and redirect methane emissions from livestock farming.

Within the digestive system of ruminants, a consortium of microorganisms thrives in the first compartment of their specialized stomach, the rumen. These microorganisms facilitate the breakdown of the complex carbohydrates through the enteric fermentation process. Methanogenic microbes harness the hydrogen produced during the fermentation process to reduce carbon dioxide, which eventually yields methane as a by-product. The process of rumination, where ruminants regurgitate cud from the rumen to the mouth, indicates a potential association between the oral and rumen microbiomes and the methane emission phenotype. Several studies have successfully predicted methane emissions from ruminant oral microbiomes. Recognizing the pivotal role of the oral and rumen microbiome's composition and activity on methane yield, our study hypothesized the potential mitigation of ruminal methane emissions by strategically selecting animals with a microbiome demonstrating a reduced propensity to produce methane. We conducted a comparative analysis of microbiome signatures and activity profiles between oral and rumen samples, exploring their correlation with the methane emission phenotype. This innovative approach proposes to infer low methane-emitting cattle via a non-invasive oral microbiome and an integration into existing breeding programs, maximizing the benefits of selective breeding.

From FASTQ to Fork – Omics Assisted Product Development

Combination of capture methyl seq technology and molecular tools to highlight imprinted loci in pig

Mathilde Perret[1], Katia Fève[1], Nathalie Iannuccelli[1], Jean-Noël Hubert[1], Juliette Riquet[1], Julie Demars[1]

1 GenPhySE, INRAE, Université de Toulouse, INRAE, ENVT, Castanet Tolosan, France

Genomic imprinting (GI) is an epigenetic phenomenon in which genes are mono allelically expressed depending on the parental origin. GI is regulated through epigenetic marks that also show parent-of-origin (PofO) methylation resulting in differentially methylated regions (DMRs) between both parental alleles. Identifying such specific patterns requires to integrate the most suitable molecular and computational tools which can be challenging especially in livestock species in which knowledge on GI is still sparse.

Based on a method developed by our team, we propose to characterize in pigs, the imprinting status of regions that are known to be imprinted in humans and mice.

We identified 135 regions in blood showing molecular signatures of GI from capture methyl-seq trio data, combining DMR information and informative variants. The regions of interest include the GNAS locus which is associated with growth traits in pigs. We showed that the GNAS region carries PofO methylation within 5 DMRs including a potentially new one. Using independent methods of genotyping and methylation sequencing, these results were validated in blood and confirmed in brain and muscle.

Our results support the relevance of our novel technology to advance the characterization of GI mechanisms especially DMRs, in pigs. We are currently working on the specificity of the porcine GNAS locus, showing the potential interest of enlarging the study of GI to a broad range of species to better understand the contribution of GI in complex traits variability. Indeed, given the crucial role played by imprinted genes in fetal and early life for growth and development but also with behavior troubles in adulthood, taking into consideration the importance of GI into the determinism of traits of major agronomic importance may help to sustainable livestock systems.

Diving deep into the full spectrum of genetic variation by pangenome visualization to facilitate trait discovery.

Joerg Hagmann[1], Christian Kubica[1], Tatsuki Tsujimoto[1], Josie Bretz[1], Sebastian J. Schultheiss[1]

1 Computomics GmbH, Tübingen, Baden-Württemberg, Germany

The commonly high genetic diversity between plant genomes of the same species and advances in sequencing technology propel the generation of whole genome assemblies, whose alignments can be efficiently summarized today in pangenome graphs. Rather than focusing on single nucleotide exchanges relative to a reference genome, comparing multiple genome sequences opens up the view on the full spectrum of genetic variation in a population, including non-reference sequence and large structural variation. This can lead to an extended set of markers to better explain and predict traits.

Possibilities to explore such novel data sets without having strong computational skills are still lacking. We are presenting Pantograph, our interactive browser of pangenome graphs capable of visualizing small- and large-scale genetic variation from single nucleotide up to whole chromosome level views in a single framework with typical browser functionalities. We will showcase the application of Pantograph on pangenomes from different species. User defined meta data like phenotypic values can be used to sort the sequences to help link larger scale genetic variation to phenotypic traits and to inspect QTL regions. The functional impact of variants can be assessed by the simultaneous display of multiple genome annotations as well as the integration of extensive gene expression and functional annotation data. With integrating public pangenome and widely used marker and whole-genome sequencing data, the goal of Pantograph is to become a multi-omics trait knowledge resource and co-pilot to facilitate trait discovery, marker development, and gene editing target detection.

From lateral root to functional nodule: Spatiotemporal transcriptomics to understand and engineer organogenesis in barley

Min-Yao Jhu[1], Giles Oldroyd[1]

1 Crop Science Centre, Department of Plant Sciences, University of Cambridge, Cambridge, United Kingdom

Increasing crop production is required to nourish the projected global population by 2050 for future food security. Nitrogen availability often constrains crop yields, and the sustainability of current food production systems is further compromised by the extensive use of chemically fixed nitrogen fertilizers. Addressing this issue requires the exploration of alternatives that empower farmers to enhance productivity sustainably. Scientists are actively investigating the potential of harnessing biological nitrogen fixation through the legume-rhizobium symbiosis for crop engineering. Building upon our current understanding of root nodule symbiosis, we aim to produce self-fertilizing crops by transferring this symbiotic relationship with nitrogen-fixing rhizobia from legumes to cereals. Establishing a functional symbiotic relationship between rhizobia and plants necessitates tightly regulated nodule organogenesis. While significant progress has been made in comprehending root nodule symbiosis in legumes over the past decade, achieving spatiotemporal resolution over the mechanisms involved in lateral root organogenesis in target crops is crucial for transferring this trait to non-nodulating species. Our research projects focus on investigating existing developmental mechanisms in Medicago nodules and barley lateral roots using spatiotemporally resolved transcriptomics to identify developmental-stage-specific and tissue-specific native promoters. Leveraging this knowledge, we aim to conduct precise spatiotemporal engineering on barley roots, regulating lateral root organogenesis to facilitate the formation of nodule-like structures capable of accommodating nitrogen-fixing bacteria.

Global Opportunities and Challenges for Agriculture

Decoding the genetic mechanisms on African swine fever tolerance: Implication for improving disease control in susceptible pig populations

Edward Okoth[1], Wen Feng[2,3], Eunice Machuka[1], Samuel Oyola[1]

1 International Livestock Research Institute, Animal and Human Health Department, Nairobi, Kenya
2 China Agricultural University, College of Animal Science and Technology, Beijing, China
3 Shenzhen Kingsino Technology Company Limited, Shenzhen, China

African Swine Fever (ASF) is one of the most important viral diseases affecting swine, with serious socioeconomic impact. The disease was first reported in Kenya in 1910 and is currently endemic in many African countries. It was reported in domestic pigs in China in 2018 and continues to spread globally. This highly contagious infection with no effective and commercially available vaccine has reached pandemic proportions, substantially impacting pork production, protein availability, livelihoods in affected countries and multifaceted aspects of local and international trade. Measures taken by affected countries have not been sufficient to stop spread. We discuss studies on genome evolutionary process of local pig breeds and the wild pig species in Africa to detect genes related to the disease. We show evidence of positive selection on the peptide binding regions of Swine Leucocyte Antigen (SLA-1, SLA-2 and DQB1) genes in domestic and wild pigs that can be targeted for association studies of ASF resistance or susceptibility traits. Complete genome assembly of the domestic and wild pigs decode the genetic mechanism of ASF tolerance associated with over-expression of siRNA of lactate dehydrogenase B (LDHB) and TRIM family genes in warthog genome that are potentially responsible for tolerance in the species. We discuss how functional validation of these and other genes or mutations identified in our studies provide evidence for association with immune and antivirus pathways. These can be targeted in molecular breeding strategies to improve ASF tolerance in pig populations and, in turn, support ASF control.

Development of AgriSeq™ targeted sequencing catalog panels for specific field and vegetable crops.

Tiffany Wang[1], Grant Poole[1], Krishna Reddy[1], Sandesh Shrestha[1], Claudio Carrasco[1], Stephane Daly[1], Weizhong Li[1]

1Thermo Fisher Scientific, Austin, TX
2Thermo Fisher Scientific, Lissieu, France
3Thermo Fisher Scientific, Carlsbad, CA

Over the years our Team at Thermo Fisher Scientific has designed over 300 custom panels for over 70 crops for customers for use with the AgriSeq™ targeted genotyping by sequencing (T-GBS) workflow on the Ion Torrent® platform. Here, we report the recent design of ready to use catalog crop panels for wheat, barley, maize, soybean, strawberry, tomato, pepper and lettuce. These panels have various marker densities for use in breeding programs for utilization in Marker Assisted Selection, Genomic Selection and screening of larger 5k panels for SNP performance evaluation and selection in specific breeding germplasm. Designs were completed for the major crops based on publicly available SNP content as well as relevant public content from our Catalog Axiom microarray platform utilizing the most up to date reference genome information at NCBI (Ruff et al. 2020, Ganal et al. 2011, Hardigan et al. 2021, Hulse-Kemp et al. 2016; Sims et al. 2012). SNP content was specifically chosen on the basis of minor allele frequency (greater than 5% where applicable), expert breeder review and public information from population screening and selection of SNPs of interest. Most of the crop panels include 20-200+ trait-based SNP marker designs for the most important economic traits of interest such as disease resistance, agronomic traits, crop quality, and yield component parameters. Results from initial screening with these panels show adequate performance for quickly screening breeding material for routine screening or to select the best performing SNP content for potential custom designs tailored to specific breeding programs. These sequencing panels can be used for routine screening of economic traits of interest or utilization in GS modelling to improve speed, accuracy and efficiency of selection in crop breeding programs. These panels represent the first of several commercial crop based targeted sequencing panels to be produced for future crops in demand.

Keywords: Catalog Crop Panels, MAS,Trait, AgriSeq, NGS

Harnessing natural genetic variation and local adaptation in sorghum for broad-spectrum resistance to the purple witchweed (Striga hermonthica)

Joel Masanga[1], Terry Felderhoff[2], Fanna Maina[3], Brian Rice[4], Clara Cruet-Burgos[5], Ghislain Kanfany[6], Baloua Nebie[7], Steven Runo[8], Geoffrey Morris[5], Jesse Lasky[1]

1 Department of Biology, Pennsylvania State University, University Park, PA
2 Department of Agronomy, Kansas State University, Manhattan, KS
3 Institut National de la Recherche Agronomique du Niger, Niamey, Niger
4 Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE
5 Department of Soil and Crop Sciences, Colorado State University, Fort Collins, CO
6 Centre National de la Recherche Agronomique, Institut Sénégalais de Recherches Agricoles (ISRA), Bambey, Senegal
7 International Maize and Wheat Improvement Center (CIMMYT), Senegal
8 Department of Biochemistry, Microbiology and Biotechnology, Kenyatta University, Nairobi, Kenya

Cereal crop yields across sub–Saharan Africa are negatively affected by Striga hermonthica Del. Benth, a hemiparasite that infects host roots and siphons its water, and photoassimilates. Although intra-specific local adaptation and natural genetic variation offer an opportunity for crop improvement, deployment of resistance materials across infested regions may be constrained by diversity of local parasite populations. Here, we interrogated whether ecotypic variations in Striga could complicate development and deployment of resistance alleles. We first tested phenotypic diversity among Striga ecotypes collected across African ecosystems with regards to response to host-derived chemical cues that induce germination of parasite seeds (strigolactones). We found some with stronger response to orobanchol than 5-deoxystrigol – major strigolactones in sorghum– suggesting limits to the value of any allele-based resistance. Next, we evaluated whether root exudates from sorghum genotypes carrying different natural mutations in the LOW GERMINATION STIMULANT 1 (LGS1) locus display differential response to Striga ecotypes in vitro. We also analyzed whether genetically similar sorghum landraces obtained from different environments are locally adapted to Striga and found some that stimulated lower parasite germination rates. Finally, we generated near isogenic lines (NILs) by introgressing loss-of-function LGS1 allele from SRN39 (a Striga-resistant genotype that possesses the lgs1-1 variant) into elite farmer-preferred backgrounds from Niger, namely Mota Maradi, Sepon82 and MR732. We found that the NILs displayed marked resistance to Striga collected from one geographical region, evidenced by significantly lower numbers of germinated Striga seeds compared to unimproved parents. In this poster, we highlight the patterns of response of these improved materials to different Striga ecotypes and discuss the far-reaching implications of the observed ecotypic parasite variations on development and deployment of resistance alleles for broad-spectrum resistance to the noxious weed.

References:
1 Gobena D, Shimels M, Rich PJ, Ruyter-Spira C, Bouwmeester H, Kanuganti S, Mengiste T, Ejeta G, Proc Natl Acad Sci, 2017, 114(17):4471-4476.
2 Bellis ES, Kelly EA, Lorts CM, Gao H, DeLeo VL, Rouhan G, Budden A, Bhaskara GB, Hu Z, Muscarella R, Timko MP, Nebie B, Runo SM, Chilcoat ND, Juenger TE, Morris GP, de-Pamphilis CW, Lasky JR, Proc Natl Acad Sci, 2020, 117(8):4243-4251.
3 Bellis ES, McLaughlin CM, dePamphilis CW, Lasky JR, Ecography, 2021,44: 1205-1217.

Merging open and close-ended survey design for cassava trait preference research

Erin E. Farmer[1], David Brown[1], Michael A. Gore[1], Hale A. Tufan[1]

1 Plant Breeding and Genetics Section, School of Integrative Plant Science, Cornell University, Ithaca, New York 14853, USA

Research into trait preferences and variety evaluation are both critical to the adoption and dissemination of improved crop varieties. This has implications for breeding programs, allowing breeders to target growers' and consumers' needs by modifying breeding goals to match known preferences. Several studies have been conducted on cassava trait preferences and varietal rankings, though these can be expensive and time consuming. Integration of disparate and varied data collected across qualitative and quantitative methodologies may be used to synthesize historical data and inform future crop improvement decisions. Here, we investigate the potential for using natural language processing (NLP), including generative pre-trained transformer models (GPTs), to generate labels from open-ended survey questions to further inform and investigate close-ended surveys. Data were collected as part of The Cassava Monitoring Survey in Nigeria in 2015. Both open and close-ended trait preference questions were asked to the same households, allowing direct comparison of the survey questions and their responses. NLP was used to successfully extract labels from the open-ended responses, and GPT models were used to tag those observations with corresponding labels. New labels, which had not been included in the close-ended questions, were found, and moreover, almost two-thirds of the 15 most frequent labels did not directly map to close-ended labels. When tagging responses, approximately 20% additional responses were captured if unmapped open-ended labels were included, as compared to close-ended labels alone. A set of corresponding labels was determined from the extracted open-ended questions' labels and the close-ended questions' predetermined labels. Comparison of household responses to each set of questions showed that there was low correlation between preferences stated in open versus close-ended questions, with values below 0.3 for all traits. Finally, labels from both were compared to Crop Ontology terms to test the coverage and representation of respondent data.

Nanopore library preparation kit biases and microbiome analysis

Ziming Chen[1], Chian Teng Ong[1], Loan Nguyen[1], Harrison Lamb[1], Oscar Gonzalez-Recio[2], Sarah Meale[3], Elizabeth Ross[1]

1 Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, St. Lucia, Australia
2 Departamento de Mejora Genética Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, CSIC, Madrid, Spain
3 School of Agriculture and Food Sciences, The University of Queensland, Gatton, Australia

Nanopore sequencing is a third-generation sequencing technique. Two common amplification-free library preparation protocols are available for Nanopore sequencing: the ligase-based method (ligation kits); and the transposase-based method (rapid kits). Different library preparation methods can generate sequencing biases, leading to the loss of single nucleotide variants or misrepresentation of microbial composition. This study aimed to compare sequencing bias between the two kits. A distinct bias towards the recognition motif (5'-TATGA-3') of MuA in the rapid kit was identified using sequencing data from cattle. Moreover, relatively lower AT contents at the sequence terminus were seen in the ligation kits. Subsequent analyses showed that the rapid kit had strong insertion and sequencing depth preference in 40 – 70 % guanine-cytosine (GC) content regions, while those of the ligation kit had a relatively even distribution. Insertion events and sequencing coverage showed a strong positive correlation in the rapid kit (R2=0.67). Furthermore, observed bacterial profiles of microbiome samples showed significant differences between the two protocols. The ligation kit showed better assembly and meta-epigenomic analysis results, likely due to longer read lengths, but no significant difference was seen. Our study indicated that careful and consistent library preparation method selection is essential for comparative microbiome studies.

Unraveling Gene Expression Dynamics in Bioenergy Crops Under Environmental Stress Challenges

Dae Kwan Ko[1,3] and Federica Brandizzi[1,2,3]

1 MSU-DOE Plant Research Lab, Michigan State University, East Lansing, Michigan, 48824
2 Department of Plant Biology, Michigan State University, East Lansing, Michigan, 48824
3 Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, Michigan, 48824

The resilience, growth, and productivity of crops face formidable challenges from environmental stressors such as drought, heat, and salinity. These challenges are expected to intensify progressively and irreversibly in the future. Plants, to adapt and flourish in such conditions, undergo swift changes in gene expression crucial for cellular and metabolic functions. Despite the critical importance of understanding these processes for both basic research and practical field applications, a comprehensive understanding of abiotic stress-induced global gene expression changes in crops has remained elusive. In our study, we sought to fill this knowledge gap by generating an extensive time-series of transcriptome datasets in sorghum. Employing RNA-sequencing analyses under three major environmental stresses—drought, heat, and salinity—we employed a well-established gene network modeling pipeline. This approach allowed us to gain deeper insights into the dynamic changes in gene expression in response to environmental stress. By utilizing dynamic time-course data on the transcriptome changes in shoots and roots of sorghum, the fifth most crucial cereal crop globally, we were able to dissect the intricate stress-, tissue-, and phase-specific gene responses. Our streamlined gene network modeling pipeline not only enhances our understanding of gene regulatory dynamics in response to environmental stress but also provides essential resources for genetic engineering and molecular breeding aimed at enhancing crop resilience. This study contributes valuable knowledge that can be leveraged for the development of more robust and stress-tolerant crop varieties, addressing the challenges posed by changing environmental conditions.

Index

	Last Name	First Name	Organization / Company	Page(s)
I	Ibáñez-Escriche	Noelia	Universitat Politècnica de València	30
J	James	Leonora	Iowa State University	75
	Jhu	Min-Yao	Crop Science Centre, Uiversity of Cambridge	103
K	Kapoor	Muskan	Iowa State University	74
	Kayondo	Fazhir	Iowa State University	21
	Khilji	Sarem	University of Idaho	26
	Kiugu	Emily	University of Minnesota	85
	Ko	Dae Kwan	Michigan State University	110
L	Lavoie	Jennifer	Thermo Fisher Scientific	65
	Lee	Kendall	HudsonAlpha	96
	Li	Forrest	UC Davis	51
	Li	Donna	UC Davis	53
	Lohr	Chaelynne	CentralStar Cooperative	29
	Londono Gil	Marisol	University of Georgia	40
	Long	Katie	John Innes Centre	91
M	Majot	Adam	Agriplex Genomics	62
	Mamedio de Freitas	Gabriel	São Paulo University	48
	Masanga	Joel	Pennsylvania State University, USA	56, 107
	McCarthy	Fiona	University of Arizona	72
	McKay	Stephanie	University of Missouri	26, 52, 72, 81

	Last Name	First Name	Organization / Company	Page(s)
M	Moenga	Susan	International Crops Research Institute for the Semi-Arid Tropics (ICRISAT)	68
	Monis Moreno	Brenda	Luiz de Queiroz College of Agriculture - University of São Paulo	27
	Morales	Abraham	Joint Genome Institute (LBNL)	69
	Mwangi	Mary	International Potato Center	82
N	Nguyen	Loan	The University of Queensland	87, 109
	Noll	Kelsey	EpiCypher, Inc	22
O	Okoth (Abworo)	Edward	International Livestock Research Institute	105
	Olabosoye	Blessing	Iowa State University	76
	Olagunju	Temitayo	University of Idaho	81
	Ong	Chian Teng	The University of Queensland	99, 109
P	Pantha	Pramod	Louisiana State University	98
	Patel	Kruti	New England Biolabs	63
	Perret	Mathilde	INRAE GENPHYSE	101
	Petroli	Cesar	International Maize and Wheat Improvement Center (CIMMYT)	41
	Powell	Owen	The University of Queensland	23, 37

	Last Name	First Name	Organization / Company	Page(s)
R	Reith	Rachel	University of Nebraska - Lincoln	47
	Reveche	Maria Ymber	International Rice Research Institute	66
	Rice	Brian	University of Nebraska - Lincoln	44, 107
	Rios	Esteban	University of Florida	46
	Ross	Elizabeth	The University of Queensland	23, 87, 99, 109
	Rusholme-Pilcher	Rachel	Earlham Institute	70
S	Sanarana	Yandisiwe	University of Pretoria	97
	Savva	Loizos	John Innes Centre	61
	Steensma	Marije	Wageningen University & Research	25
	Stegemiller	Morgan	University of Idaho	26, 81
	Streck	Chris	n6 Tec, Inc.	95
	Stull	Caleb	University of Missouri - Columbia	73
T	Tabet	Joe Menwer	University of Georgia	42
	Trujano Chavez	Mitzilin Zuleica	University of Georgia	39
V	Van Eenennaam	Alison	UC Davis - Department of Animal Science	84
	Veroneze	Renata	Universidade Federal de Viçosa and University of Georgia	36
	Villiers	Kira	Queensland Alliance for Agriculture and Food Innovation (QAAFI), at the University of Queensland	49

AGBT2025™
AGRICULTURE

SAVE THE DATE

MARCH 31 – APRIL 2, 2025

CARIBE ROYALE RESORT | ORLANDO, FL