TOOLS FOR POLYPLOIDS

Training Workshop 2022

This project is funded by USDA NIFA Specialty Crop Research Initiative Award # 2020-51181-32156 (09/01/20 - 08/31/24)

Contact Us!

Email: toolsforpolyploids@tamu.edu

Twitter: @polyploidtools

This project is funded by USDA NIFA Specialty Crop Research Initiative Award # 2020-51181-32156 (09/01/20 - 08/31/24)

Contents

Workshop Agenda: Thursday, January 13, 2022

January 13		
Time (US Pacific)	**Presenter**	**Topic**
8:30	David Byrne	Introduction
8:45	Dorrie Main Delany Baum	Community Resource website
9:00	Peter Bourke Chris Maliepaard Roeland Voorrips	polymapR and polyqtlR updates
9:50	Jibran Tahir Cyril Brendolise	Genetic Maze for polygenic traits in heterozygous polyploid kiwifruit population
10:05	Paterne Agre et al.	Identification of QTLs Controlling Resistance to Anthracnose Disease in Water Yam (Dioscorea alata)
10:20	Break	
10:30	Marcelo Mollinari Gabriel Gesteira Cris Taniguti	MAPpoly and QTLpoly : Use, progress and case studies
11:30	Cris Taniguti et al.	Reads2Map: Practical and reproducible workflows to build linkage maps from sequencing data
11:45	Alejandro Thérèse Navarro et al.	Smooth Descent: a ploidy-agnostic algorithm to improve linkage mapping in the presence of genotyping errors
12:00	Office Hours	Main room – polymapR and polyqtlR
1:00	Jeff Endelman	New software features for GWAS
1:30	Mason Chizk et al.	Novel Genomic Resources Enable Association Mapping in Tetraploid Blackberry
1:45	Jeekin Lau Cris Taniguti Oscar Riera-Lizarazu David Byrne	Comparison of polymapR/polyqtlR and MAPpoly/QTLpoly in a tetraploid rose dataset
2:20	Break	
2:25	David Byrne	Announcements
2:30		"Office Hours" Breakout room #1 – MAPpoly/QTLpoly Breakout room #2 – GWASpoly Breakout Room #3 – Software comparisons
3:00	Poster session	Find breakout room assignments here. These are accessible through the Main room.
4:00	Adjourn	

Workshop Agenda: Friday, January 14, 2022

January 14		
Time (US Pacific)	**Presenter**	**Topic**
8:30	David Byrne	Welcome and announcements
8:40	Roeland Voorrips Chris Maliepaard Peter Bourke	Simulation Tools: PedigreeSim and GenoSim
9:45	Ranjana Bhattacharjee et al.	Harnessing the promising fruits of genomics-assisted breeding for yam improvement
10:00	Break	
10:10	Tim Millar Susan Thomson	Haplotype assembly using sequencing data
11:10	Jeff Endelman	Overview of genomic selection, Stagewise
12:00	Office Hours	Main room – Simulation Tools
1:00	Jeff Endelman	Genomic selection: potato case study
1:30	Patricio Muñoz	Genomic selection: blueberry case study
2:00	Olivia Angelin-Bonnet et al.	Investigating genotype-phenotype relationships for tuber bruising in autotetraploid potatoes
2:15	Nweze Peter et al.	Evaluation of *Musa* Accessions Indigenous to Benin Republic for Resistance to Black Sigatoka Disease
2:30	Joao Vitor Nomura et al.	Genome-Wide Association for Drought Tolerance in Potato
2:45	David Byrne	Final comments
3:00		"Office Hours" Breakout room #1 – Haplotype Assembly Breakout room #2 – Stagewise
4:00	Adjourn	

Major Presentations

Community Resource Website
Dorrie Main and Delany Baum

This presentation will include a tour of the website and a review of the project's social media pages.

polymapR and polyqtlR
Peter Bourke, Chris Maliepaard, and Roeland Voorrips

In this workshop, participants will be re-acquainted with Wageningen's polymapR / polyqtlR pipeline for linkage map construction and QTL analysis in polyploid F1 populations. Demonstration of new or extra software features not previously covered will be highlighted. These include using probabilistic genotypes in place of discrete genotypes, automatic building of multi-QTL models via co-factor analysis, visualizing meiotic pairing behavior and genotype curation (spotting and correcting genotyping errors).

MAPpoly and QTLpoly
Marcelo Mollinari, Gabriel Gesteira, and Cris Taniguti

In this workshop, we will show how to construct genetic maps and perform QTL analysis in multiparental polyploid populations using new versions of the R packages MAPpoly and QTLpoly. In the hands-on section, we will use a tetraploid dataset to show the software features and present results in three interconnected hexaploid sweetpotato populations. We will also introduce VIEWpoly, an R package for visualizing and exploring results from polyploid computational tools using an interactive graphical user interface.

New Software Features for GWAS in Polyploids
Jeffrey Endelman

This presentation will cover new features added to GWASpoly in 2021, including (1) partial R2 values, (2) mixed ploidy datasets, and (3) VCF files as input.

Software comparisons
Jeekin Lau, Cris Taniguti, Oscar Riera-Lizarazu, and David Byrne
This presentation will discuss a comparison on the performance and usability of linkage mapping/QTL software on empirical tetraploid rose dataset. Software compared are MAPpoly/QTLpoly, polymapR/polyqtlR, PolyOrigin/diaQTL. Linkage map quality comparisons and QTL comparisons of the different software. Comparison on general ease of use and computational demands of each software will also be noted.

Simulation Tools Workshop: PedigreeSim and GenoSim
Chris Maliepaard, Roeland Voorrips, and Peter Bourke

The simulation software we present generates data similar to data observed in actual experiments, but where the true genetic make-up of individuals is known. This is useful for developing and testing software for genetic analysis, and can help to choose the best tools for a give data set. It also allows to compare different breeding strategies.

Haplotype assembly
Tim Millar and Susan Thomson

We have developed a new software package "MCHap" for assembly of micro-haplotypes in polyploids using Markov chain Monte Carlo simulation. This software is capable of assembling known multi-allelic SNPs into haplotypes within pre-defined genomic loci. We will walk through the functionality of the tool and show worked examples of capture-sequence genotype datasets.

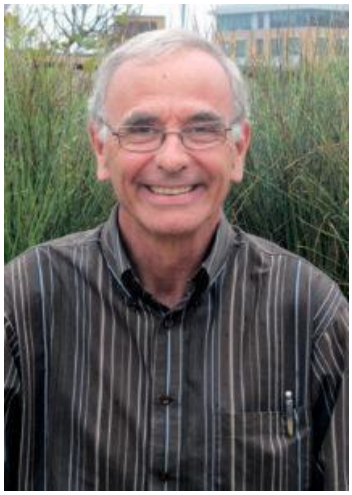
Genomic selection: Stagewise
Jeffrey Endelman
This presentation discusses the theory and principles of genomic selection, including BLUP, reliability, correlated traits, and two-stage analysis of multi-environment trials.

Potato case study
Jeffrey Endelman
Several workflows in R package StageWise will be illustrated using potato breeding data.

Blueberry case study
Patricio Muñoz

In this talk we will cover the operational application of genomic selection (GS) to blueberry (Vaccinium corymbosum). Blueberry is an autotetraploid, highly heterozygous species with high level of inbreeding depression. Thus, we will also discuss general challenges encounter when applying GS in autotetraploid species. Finally, we will cover the process of optimization to make GS operationally possible in the University of Florida blueberry breeding program.

Presenter Information



Dr. David Byrne (PD), the Basye Endowed Chair in Rose Genetics, and a Professor in Horticultural Sciences at Texas A&M University. He works internationally in Prunus and Rosa breeding and genetics. He led the SCRI Combating Rose Rosette Disease project, worked with both SCRI RosBREED projects and currently chairs the National Clean Plant Network for Roses. He routinely works with numerous scientists, private breeding programs, ornamental nurseries, marketers, and producers as well as ornamental hobbyist groups. He has, in collaboration with Dr. Klein, and Dr. Riera-Lizarazu is working to incorporate genomic approaches (MAB, GWAS, GS) into the rose breeding program to enhance disease resistance (RRD, black spot, Cercospora), flower productivity, and horticultural quality of roses. Over the past decade, he has mentored 18 graduate students, published 37 refereed and 27 non refereed publications, has given about 140 talks, and released 23 Prunus and 6 Rosa cultivars. He is the Project Director and coordinates the training aspects of the project.

Dr. Dorrie Main is a Professor of Bioinformatics in the Department of Horticulture at Washington State University. She runs a highly collaborative, multi-disciplinary research program that has been funded through public (NSF, USDA) and industry (cotton, tree fruit, and legume) sources. The research develops database resources for specific crops (e.g., Genome Database for Rosaceae, CottonGen, Citrus Genome Database, Cool Season Food Legume Database, Genome Database for Vaccinium), online computational tools (e.g. GenSAS), sequence analysis pipelines (e.g. RefTrans), and generic database platforms (e.g. Tripal). The lab also identifies genomic regions and markers controlling important traits and studies genome. These efforts seek to provide genomic, genetic, and breeding resources to enable basic, translational and applied research. They host the AgBioData website.

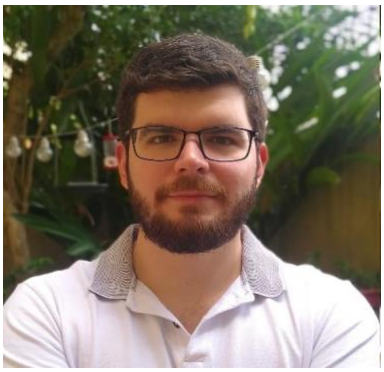
Delany Baum is the project coordinator for the Tools for Polyploids project. She maintains the project's website and manages the project's social media pages. Delany received her Master of Science degree in environmental science with an emphasis in education from the University of North Texas in 2020 and has conducted research in the fields of ecology and botany.

Originally with a background in physics and pure mathematics (UCC Cork, Ireland), **Dr. Peter Bourke** has also trained in horticulture and permaculture design. He worked for over four years in vegetable seed production and genetic resource conservation with the Irish Seed Savers Association, Ireland. Peter gained a Master of Science degree in Plant Sciences from Wageningen University & Research in the Netherlands and continued with doctoral research into polyploid quantitative genetics and mapping under the guidance of Prof. Dr. Richard G.F. Visser, Dr. Chris Maliepaard, and Dr. Roeland E. Voorrips. He is currently a researcher and lecturer in the Plant Breeding group in Wageningen.

Dr. Chris Maliepaard joined Wageningen University – Plant Breeding in 2007. He is currently an associate professor and research group leader of on the 'Quantitative Aspects of Plant Breeding' group. He carries out research projects about mapping, QTL analysis, and marker-assisted selection, mostly related to polyploid crops; other research focuses on the prediction of phenotypic traits from ~omics data sets. His expertise is in plant breeding, quantitative genetics, statistics, and quantitative analysis aspects of systems biology and bioinformatics. He teaches about these topics in several MSc, Ph.D., and international courses.

Dr. Roeland Voorrips is a senior researcher at Wageningen Research – Plant Breeding Before joining Wageningen Research in 1986 he worked at a vegetable seed company.. He has studied disease and pest resistance in many vegetable crops, and he has been working on quantitative and population genetics methods, including linkage and QTL mapping. In 2010, he became involved in the development of genetic tools for polyploid crops. He has published several genetics software packages, including MapChart, Pedimap, PediHaplotyper, fitTetra / fitPoly, and PedigreeSim, and contributed to others including joinmap and polymapR.

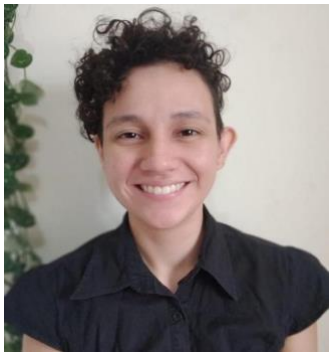
Dr. Marcelo Mollinari is a Senior Research Scholar at North Carolina State University working with the genetics of complex polyploids. He is the creator and maintainer of the R package MAPpoly. and co-created OneMap, an R package for linkage mapping in outcrossing species. In 2008, Marcelo worked at the University of São Paulo, Brazil where he assisted the development of the software SuperMASSA. In 2014, he was a Visiting Researcher at Purdue University and in 2016, he started to work on the Genomic Tools for Sweetpotato Improvement project. He, he implemented MAPpoly and built three ultra-dense multilocus genetic maps of biparental hexaploidy sweetpotato populations with the reconstruction of the full haplotypes of the parents and offspring. He also consults with researchers from several other polyploid species, such as potato, forage crops, kiwifruit, and strawberry. To check Marcelo's work, computer codes, datasets, and presentations, please visit his GitHub page at https://github.com/mmollina.



Dr. Gabriel Gesteira is a postdoctoral research scientist at North Carolina State University (NCSU). He has a Ph.D. in Plant Breeding and Genetics from the "Luiz de Queiroz" College of Agriculture, University of São Paulo (ESALQ/USP), and is the current maintainer of QTLpoly, an R package for QTL mapping in autopolyploids. His current work focuses on the development of computational tools and their application for linkage analysis, QTL mapping, and genomic prediction in polyploid species. Gabriel's interests involve polyploidy, plant breeding, quantitative genetics, statistical genetics, and programming. He is part of the SCRI Polyploid Computing Support Group.

Dr. Jeffrey Endelman is an Associate Professor at the University of Wisconsin–Madison and leads the university potato breeding program. He is a co-developer of 16 potato varieties and several software packages for genomics-assisted breeding, including *rrBLUP*, *GWASpoly*, *Stagewise*, and *diaQTL*. Endelman has co-authored 35 publications and served on the graduate thesis committees of 26 students. He is a member of the editorial boards for *Genetics*, *Theoretical and Applied Genetics*, and *The Plant Genome*.

Dr. Jeekin Lau is a postdoctoral research associate at Texas A&M University. Jeekin currently works on utilizing and comparing available genetic computational tools for tetraploid roses. He has experience using tetraploid SNP-array-based genotyping, genotype calling, linkage mapping, and QTL mapping for disease and horticultural traits. Jeekin is part of the SCRI Polyploid Computing Support Group.



Dr. Cris Taniguti is a post-doctoral research associate in the Department of Horticultural Sciences at Texas A&M University. Her current work focuses on the development of bioinformatics workflows to perform SNP and dosage identifications comparing software efficiency. She developed the VIEWpoly app, a graphical environment to evaluate linkage map and QTL analysis results. She is the current maintainer of OneMap, an R package to build linkage maps in a F_1 diploid outcrossing population. Cris is part of the SCRI Polyploid Computing Support Group.

Dr. Maria Caraza-Harter is a post-doctoral research associate at the University of Wisconsin-Madison in the Endelman Lab. She earned a Ph.D. degree in Plant Breeding and Plant Genetics at UW-Madison studying the genetics of skin set and color in red potatoes. Her main interests are in quantitative genetics and genomic association for quality traits in tetraploid potatoes. Maria is part of the SCRI Polyploid Computing Support Group.

Dr. Oscar Riera-Lizarazu, Co-Director: Dr. Riera-Lizarazu is an Associate Professor in the Department of Horticultural Sciences at Texas A&M University, He works on rose genetics and breeding with the goal of developing, testing, and releasing improved varieties of roses with regional and national adaptation as well as conducting research on the use of genomics-based tools for rose variety development and understanding the genetic basis of traits in *Rosa* and related horticultural crops. Dr. Riera-Lizarazu has over 20 years of national and international research experience in plant genetics, cytogenetics, and breeding. He has held senior-level positions in the private sector at Dow AgroSciences and Corteva Agriscience at the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) in India. .

Tim Millar is a member of the Bioinformatics team at Plant and Food Research and a Ph.D. student at Otago University, New Zealand. His interests include statistical genetics, Bayesian statistics, and software development. His Ph.D. topic is on identifying selective sweeps in a mixed-ploidy Kiwifruit (*Actinidia* spp.) breeding program.

Dr. Susan Thomson is a senior Bioinformatician and researcher at The New Zealand Institute for Plant & Food Research (PFR). Susan started her plant research life as a member of the PFR team involved in the International Potato Genome Sequencing Consortium that published the potato genome and high-density linkage map. She then moved into the field of Bioinformatics to continue the analysis of genomic and genetic data from several polyploid crop species including potato, kiwifruit, and blueberry. She is now part of a team that is focused on the application of sequence-based genotyping and haplotyping to polyploid breeding material as well as investigating the genomics and genetics of inter-specific polyploid hybrids.

Dr. Patricio Muñoz is an associate professor leading the blueberry breeding and genomics program at the University of Florida Horticultural Science Department www.blueberrybreeding.com. His research consists in the improvement of cultivars with a local focus but a global impact, and in developing and implementing genomic tools to accelerate the breeding process with a focus on outcrossing polyploid species. He has released or collaborated for the release of 10 blueberry cultivars, 4 forage cultivars, and 7 turfgrass cultivars. He also led the development of one phone App for helping growers with scouting blueberry pests and diseases, as well as to obtain details of blueberry cultivars (UF/IFAS Blueberry Growers App). Currently, he is working to develop low chilling requirement blueberries with better flavor, sensorial and nutraceutical traits to enhance the consumption of blueberries while still working on traditional crop traits that help producers stay competitive.

Oral Presentation Abstracts

Genetic Maze for polygenic traits in heterozygous polyploid kiwifruit population

Jibran Tahir and Cyril Brendolise. The New Zealand Institute of Plant and Food Research.

Polyploidy is a key driver of significant evolutionary changes in plant species. The genus Actinidia (kiwifruit) exhibits multiple ploidy levels, which contribute to novel fruit traits, high yields, and resistance to the pathogen. Breeding programs focus on tetraploid cultivars, because of their resilience to the Psa disease, as well as robust fruit quality traits and size. However, there remains a substantial knowledge gap concerning the chromosomal biology and complex gene-trait associations in polyploid kiwifruit, which also exhibits high heterozygosity due to dioecy, compared with diploid species. Here we present the first case study where we employed CaptureSeq for genotyping 235 F1 heterozygous individuals of a tetraploid kiwifruit mapping population using 10K baits which generated a total of 725,175 raw SNP variants. Using this dataset we first identified caveats in the population structure by studying kinship and recombination landscape. We then performed dosage counts, using an empirical Bayesian analysis (Updog, “normal model”), and selected 45,436 high-quality SNPs across 188 F1s for generating ultra-dense linkage maps for 29 chromosomes, using polymapR. To further assess the quality of our maps, we estimated the Identity-by-descent (IBD) probabilities. We also tested the nature of pairing among chromosomes and provided the first

evidence from a mapping population that the inheritance in polyploid kiwifruit species is mostly polysomic with preferential pairing in a few LGs. Finally, we performed QTL mapping for two key polygenic traits for breeding (bacterial resistance and flowering) and a monogenic trait for breeding operation (sex) and calculated the additive effects from various QTLs for a trait. We also tested an HRM-based SNP marker workflow downstream of QTL mapping, to assess the possibility of marker-assisted selection for these traits. Our study provides a detailed view of multiple challenges and successes in solving the genetic maze of polygenic traits in a polyploid dioecious species.

Identification of QTLs Controlling Resistance to Anthracnose Disease in Water Yam (*Dioscorea alata)*

Paterne Agre, Kumar Lava, Robert Asiedu, Asrat Asfaw. International Institute of Tropical Agriculture, Ibadan, Nigeria.

Anthracnose disease, caused by the fungus *Colletotrichum gloeosporioides,* is the primary cause of yield loss in water yam (*Dioscorea alata*), the widely cultivated species of yam. Resistance to yam anthracnose disease (YAD) is a prime target in breeding initiatives to develop cultivars with durable resistance for the sustainable management of the disease in water yam cultivation. This study aimed at tagging quantitative trait loci (QTL) for anthracnose disease resistance in a bi-parental mapping population of D. alata. Parent genotypes and their recombinant progenies were genotyped using the GBS platform and phenotyped in two crop cycles for two years. A high-density genetic linkage map was built with 3,184 polymorphic SNP markers well distributed across the genome covering 1460.94 cm total length. On average 163 SNP markers were mapped per chromosome with 0.58 genetic distance between SNPs. Four QTL regions related to yam anthracnose disease resistance were identified on 3 chromosomes. The proportion of phenotypic variance explained by these QTLs ranged from 29.54 to 39.40%. The QTL regions identified showed genes that code for known plant defense responses such as GDSL-like Lipase/Acylhydrolase, Protein kinase domain, and F-box protein. The results from the present study provide valuable insight into the genetic architecture of anthracnose resistance in water yam. The candidate markers identified herewith form a relevant resource to apply marker-assisted selection as an alternative to a conventional labor-intensive screening for anthracnose resistance in water yam.

Reads2Map: Practical and reproducible workflows to build linkage maps from sequencing data

Cristiane H. Taniguti[1], Lucas M. Taniguti[3], Gabriel S. Gesteira[2], Thiago P. Oliveira[3], Jeekin Lau[1], Getulio C. Ferreira[3], Rodrigo R. Amadeu[3], David Byrne[1], Oscar Riera-Lizarazu[1], Guilherme S. Pereira[2], Marcelo Mollinari[2], and Augusto F. Garcia[3]. [1]Texas A&M University, College Station, TX. [2]North Carolina State University, Raleigh, NC. [3]University of Sao Paulo, Sao Paulo, Brazil.

High-throughput sequencing methods produce millions of sequence reads that need to be processed by bioinformatic tools before being applied in genetics research. For each step of the procedure, such as alignment of reads, SNPs identification, and genotype calling, several tools are available, all with different methods and parameters to be selected by users. Changes in a single parameter in the pipeline can cause downstream consequences in the analysis quality. Because the genetic properties of meiotic events are well-known, it is possible to identify low-quality markers using linkage analysis. Genotyping errors lead to an overestimation of recombination events amount, inflated linkage map distances, and issues while grouping and ordering markers. Thus, good-quality genetic maps validate all upstream procedures and help to identify the best combinations of software and parameters. Here, we present the Reads2Map workflows to build linkage maps from sequencing data of experimental F1 outcrossing populations testing combinations of upstream tools. The workflows are written with Workflow Description Language (WDL) that offers a comprehensive structure and metadata for each step, making it easier for users to adapt specific parameters. WDL also allows interface with containers to increase reproducibility, facilitate access to diverse software, and use in high-performance computing or cloud service environments. The final workflow output is the input for the Reads2MapApp, a Shiny app, which allows interactive visualization of the produced genetic maps and selection of the best pipeline. We demonstrate Reads2Map workflows and Reads2MapApp using both simulated and empirical RADseq data.

Smooth Descent: a ploidy-agnostic algorithm to improve linkage mapping in the presence of genotyping errors

Alejandro Thérèse Navarro, Peter Bourke, Eric van de Weg, Paul Arens, Richard Finkers, Chris Maliepaard.
Wageningen University and Research, Wagengingen, the Netherlands.

Linkage mapping is an approach to order markers based on recombination events. Mapping algorithms cannot easily handle genotyping errors, which are common in high-throughput genotyping data. To solve this issue, strategies have been developed, aimed mostly at identifying and eliminating spurious genotypes. One such strategy is SMOOTH (van Os et al. 2005), an iterative algorithm to detect genotyping errors. Unlike other approaches, SMOOTH can also be used to impute the most probable alternative genotypes, but its application is limited to diploid species and to markers heterozygous only in one of the parents. We adapted SMOOTH to expand its use to any marker type and to autopolyploids with the use of identity-by-descent probabilities, naming the updated algorithm Smooth Descent (SD). We applied SD to real and simulated data, showing that in the presence of genotyping errors this method produces better genetic maps in terms of marker order and map length. SD is particularly useful for error rates between 5% and 20% and when error rates are not homogeneous among markers or individuals. Moreover, the simplicity of the algorithm allows hundreds of thousands of markers to be efficiently processed, thus being particularly useful for error detection in high-throughput data. We implemented SD within an R package, SmoothDescent that can perform error detection, genotype imputation, and iterative mapping for diploids and autopolyploids.

Novel Genomic Resources Enable Association Mapping in Tetraploid Blackberry

Mason Chizk[1], Margaret Worthington[1], Carmen Johns[1], Carly Godwin[1], John R. Clark[1], Rishi Aryal[2], and Hamid Ashrafi[2]. [1]University of Arkansas, Fayetteville, AR. [2]North Carolina State University, Raleigh, NC.

Until recently, large-scale association mapping studies in autopolyploid crop species have been impractical due to cost of sequencing with adequate read-depth, computing limitations, and software availability. In a collaborative effort that leverages targeted next generation sequencing (NGS) technologies and newly available software packages to overcome these obstacles, we report the first genome wide association study (GWAS) in autotetraploid blackberry (Rubus L. subgenus Rubus Watson), identifying numerous quantitative trait loci (QTL) for potential use in marker assisted selection (MAS). Using NGS via the Capture-Seq platform, 35,054 well-distributed probes were designed from the 'Hillquist' (Rubus argutus Link.) reference genome. From 2019 to 2021, a panel of 307 commercially available fresh-market blackberry cultivars and University of Arkansas Division of Agriculture breeding selections were sequenced and phenotyped for an array of traits that encompass fruit morphology, postharvest texture, acidity, sweetness, seed traits, thorniness, and plant height. Using the designed probe set, 81,150 high-quality single nucleotide polymorphisms (SNPs) were identified with greater than 150x average read-depth. Through phenotypic analyses, high broad-sense heritabilities exceeding 0.75 were observed for multiple quantitative traits including primocane height, thorn density, fruit weight, fruit length, drupelet number, 100-seed weight, and seed width/length ratio. QTL were identified for all for each of these traits except fruit weight. A QTL for fruit firmness was identified on Ra04, less than 0.5 Mb away from two polygalacturonase (PG) homologs and two pectin methylesterase (PME) homologs. Homologs for both candidate genes have been widely implicated in texture variation for other crops. A shared QTL for titratable acidity (TA) and pH was identified on Ra05 near three malate synthases, two vacuolar malate transporters, one MYB transcription factor, and one phosphoenolpyruvate carboxylase PEPC. Other major QTL associated with primocane height and thornlessness were discovered on chromosome Ra04 at 25 and 33 Mb, respectively. These data will not only be used to design the first diagnostic markers for MAS in blackberry, but will serve as a foundational training dataset for genomic selection to achieve gains in heritable traits with few to no high-impact QTL, such as soluble solid content (SSC) and fruit weight.

Harnessing the promising fruits of genomics-assisted breeding for yam improvement

Ranjana Bhattacharjee[1], Asrat Amele[1], Jessica B. Lyons[2], Jessen Bredeson[2], Paul D. Fraser[3], Lukas Muller[4], Agre Paterne[1], P. Lava Kumar[1], Michael Abberton[1], Patrick Adebola[1], Daniel Rokhsar[1] and Robert Asiedu[1]. [1]International Institute of Tropical Agriculture, PMB 5320, Ibadan, Nigeria. [2]University of California – Berkeley, California, USA. [3]Royal Holloway University of London, London, UK. [4]Boyce Thompson Institute, Cornell University, Ithaca, USA.

Advances in next-generation sequencing (NGS), genotyping technologies coupled with high-performance computation approaches, and partnerships with advanced research institutes has paved many possibilities to apply

genomics tools to advance crop-breeding programs in several model and non-model plant species. This enabled the application of 'genomics-assisted breeding' or 'whole genome selection'. Breeding for yam improvement targets highly diverse biotic and abiotic constraints, whilst meeting complex end-user quality preferences to improve livelihoods of beneficiaries in developing countries. Achieving breeding targets and increasing the rate of genetic gains in yam, with long breeding cycles, and genomes with high heterozygosity and different ploidy levels, is challenging. However, a variety of genomic approaches has already been incorporated at various levels, including sequencing of draft genomes of some species, and application of NGS to genotype genetic resources (genebank accessions) and breeding lines for further exploitation using genomic approaches, such as genetic diversity, linkage mapping and QTL (quantitative trait loci) analysis while pipelines are established to test GWAS (genome-wide association study) and GS (genomic selection). The use of GS will become more evident in yams as the tools become more efficient for polyploid genotypes, incorporating dominance and epistatic effects, as well as multi-location environmental effects, and progeny selection. This will need to be coupled with high-throughput phenotyping, knowledge from other -omics approaches (e.g., gene expression via transcriptomics, protein function via proteomics, and metabolic pathways via metabolomics), and big data platform to allow the identification of molecular markers linked to complex traits, the dissection of genetic variability, identification of putative candidate genes, and their causative alleles for gene expression or gene function. The yam program can expand and benefit by bringing together a diverse range of disciplines and partners, and knowledge sharing from other root and tuber crops. The multi-species nature of the crop provides a huge platform to even learn from different species and expand.

Investigating genotype-phenotype relationships for tuber bruising in autotetraploid potatoes

Olivia Angelin-Bonnet[1], Susan Thomson[2], Patrick J. Biggs[1], Matthieu Vignes[1], and Samantha Baldwin[2]. [1]Massey University, Palmerston North, NZ. [2]Plant and Food Reserach, Lincoln, NZ.

Tuber bruising of tetraploid potato is an important quality trait as it affects the appearance and flavour of the tubers and thus impacts their fitness for sale. The development of potato lines that are more resistant to bruising is therefore a desirable objective for breeding programs, rendering the genetic analysis of this trait an important task. In this study, we investigated the biological mechanisms underlying tetraploid potato tuber bruising using multi-omics data. Genotype by sequencing using exon capture obtained from a breeding population of half-sibling families was used to uncover regions of interest for the bruising phenotype, as well as other agronomic traits of interest. In addition, we employed a Systems Biology approach to obtain a more holistic and comprehensive view of the molecular mechanisms involved in tuber bruising and bridge the gap between genetic variations and phenotype. To this end, RNA sequencing and metabolomics data were also obtained, and GWAS, differential expression analysis, and multi-omics data integration were leveraged in order to detect molecular features (i.e. genomic variants, genes, and metabolites) involved in tuber bruising. We demonstrate that even as capture sequencing only allows us to measure genetic variations in a subset of the genome, it is possible to uncover interesting and biologically meaningful genotype-phenotype associations, especially when combining the GWAS results with other omics datasets. Moreover, these associations were obtained with samples selected from a breeding program, demonstrating that available data from populations not specifically designed for association study can be used to uncover genomic regions potentially associated with a trait of interest.

Evaluation of Musa Accessions Indigenous to Benin Republic for Resistance to Black Sigatoka Disease

Nweze Nwabueze Peter[1], Enoch G. Achigan-Dako[2], Celestine Azubuike Afiukwa[1], and Happiness O. Oselebe[1]. [1]Ebonyi State University, Nigeria. [2]University of Abomey Calavi, Benin Republic.

Black sigatoka disease caused by Pseudocercospora fijiensis is the most destructive disease of Musa species. The aim of this study was to identify Musa accessions cultivated in Benin Republic that possess host resistance capacity to black sigatoka disease. A total of 72 Musa accessions were used in the study including 58 local accessions and 14 accessions from International Musa Transit Centre (ITC) in Belgium. The experiment was laid out using the Augmented Split Plot Design at Misserete, in Benin Republic. Agronomic data were collected 9 months after planting according to standard evaluation method for the disease including; symptom appearance (SA, in days), rate of leaf death (%LD), youngest leaf spotted (YLS), number of standing leaves (NSL), index of non-spotted leaf (INSL), area under disease progress curve (AUDPC) and disease severity index (DSI). The presence of genes for

black sigatoka resistance (bs1) was evaluated using a gene specific SSR marker and one unspecified disease resistant gene (UDRG) in Musa. DNA sequencing was done using ABI sequencer. The results revealed high significant difference ($p<0.0001$) in sigatoka disease incidence and severity among the accessions with the rate of symptom appearance (SA) ranging from 10 to 36 days, NSL and YLS ranged from 3.0 to 14.0 and 2.0 to 6.0 respectively, while the values of INSL, DSI and AUDPC ranged from 12.5 to 60.0, 13.33 to 52.0 and 58.0 to 239.0, respectively. A dendogram based on disease sensitivity data showed 3 distinct clusters (A-C). cluster A contains 20 resistant accessions, cluster B is made up of 16 susceptible accessions, while cluster C includes 36 intermediate resistant accessions. Also, the dendogram of the amplified resistance gene sequences and the gel matrix also clustered the Musa accessions into 3 groups which showed a little variation with from their morphological pattern of resistance to the disease. Multiple sequence alignment result based on UDRG sequences showed missing nucleotide sequence in some susceptible accession. The marker-trait association analysis revealed a mutual occurrence between UDRG_300 and the morphological indicators of resistance to the disease (YSL, NSL and INSL). The study therefore identified black sigatoka resistant Musa accessions among Benin Republic cultivars and provided a genetic tool (UDRG_300) which can be exploited for improvement of the crop for sigatoka disease resistance.

Genome-Wide Association for Drought Tolerance in Potato

Joao Nomura[1], Moctar Kante[2], Elisa Salas[2], Hannele Lindqvist-Kreuze[2], Fernando Angelo Piotto[1], and Jeffrey Endelman[3]. [1]Escola Superior de Agricultura Luiz de Queiroz, Sao Paulo, Brazil. [2]International Potato Center, Lima, Peru. [3]University of Wisconsin - Madison, Madison, WI.

Potato is one of the most important crops for humans, especially in developing countries. Extreme climate events are becoming more frequent, and drought is one of the abiotic stresses that can severely impact yield. To understand the genetic response for drought tolerance in potato, 655 clones were evaluated at two semi-arid, lowland sites in Peru. All clones were evaluated under full and deficit irrigation at both sites. Water reduction in the deficit treatment began when 5% of the plants started to flower, which is a morphological sign of tuber initiation. The interval between irrigation events in the full treatment was 2-3 days vs. 15 days in the deficit treatment. Targeted amplicon sequencing (DArTag) was used for genotyping, which generated 2244 SNPs after quality control. Genotype calls were made using the R/updog package, and quantitative genetic analyses used the R/StageWise package. The mean yield under full vs. deficit irrigation was 533 vs. 204 g plant-1, which represents a yield loss of 62%. Across the panel of 655 clones, the yield reduction ranged from 42 – 84%. No marker-trait association was detected for yield under full irrigation, but under deficit irrigation we observed a GWAS peak at the beginning of chromosome 5 that accounted for R2 = 21% of the breeding value. This region corresponds to the location of CDF1, a well-known gene affecting potato maturity, and which more recently has been implicated in water homeostasis. When yield under drought was plotted against the visual ratings of vine maturity, there was a tendency for earlier clones to yield more than late clones. The average yield under drought increased by 21.5 g plant-1 per allele dose at the marker linked to CDF1. Even though drought tolerance is likely affected by many genes, CDF1 appears to be an important target for genetic improvement of this trait.

Poster Presentation Abstracts

Breakout room #4.

QTL analysis of the “Tanzania” x “Beauregard” sweetpotato mapping population for resistance to *Meloidogyne enterolobii*

Simon Fraher[1], Tanner Schwarz[2], Bonny Oloka[3], Ken Pecota[1], Chris Heim[1], Adrienne Gorny[2], G. Craig Yencho[1]. [1]Department of Horticultural Science, North Carolina State University, Raleigh, NC. [2]Department of Entomology and Plant Pathology, North Carolina State University, Raleigh, NC. [3]Root Crops Program, National Agricultural Research Organization, Entebbe, Uganda.

Due to its highly heterozygous hexaploid nature, sweetpotato (I. batatas) lags behind other crops in terms of genomic tools. New tools and strategies have afforded opportunities to associate sequence data with phenotypic data through QTL analysis, especially the publication of a diploid reference genome for *I. trifida* in 2018, and open-

source software packages like QTLpoly and MAPpoly. QTL analysis is expected to be instrumental in accelerating breeding in this crop, notably for nematode resistance and nutritional factors. We analyzed the progeny of the sweetpotato biparental mapping population 'Tanzania' x 'Beauregard' (TB) representing 250 genotypes (including 4 check lines) for resistance to the emergent plant parasitic nematode, Meloidogyne enterolobii (M.e.). Parent 'Tanzania' is highly resistant, and parent 'Beauregard' is highly susceptible. Bioassays showed clear bimodal segregation for resistance, suggesting a simplex major allele conferring resistance. Using the R package QTLpoly and the I. trifida reference genome, we discovered a major QTL peak at base pair 7,039,636 (79.21cM) of linkage group 4 of I. batatas associated with resistance to M.e. This analysis suggests variability in M.e. resistance within the TB population can largely be ascribed to genetic differences amongst the progenies (h^2 =66.9%). The next steps will search for flanking markers associated with these genotypes and attempt to identify markers which can be screened in the seedling stage, reducing the need for costly and laborious bioassays with this quarantined pest.

Breakout room #5.

Octoploid Strawberry Linkage Map from Reduce Representation Sequencing SNP Markers

Jose Guillermo Chacon[1], *Marcelo Mollinari*[1], *Bode A. Olukolu*[2], *Zhao-Bang Zeng*[3], *Gina E. Fernandez*[3]. [1]Department of Horticultural Science, North Carolina State University, Raleigh, NC. [2]Department of Entomology and Plant Physiology, University of Tennessee, Knoxville, TN. [3]Bioinformatics Research Center, North Carolina State University, Raleigh, NC.

The cultivated strawberry (Fragaria ×ananassa L.) is an allopolyploid (2n= 8x = 28) with a complex genomic composition that hindered genetic and genomic studies, as the similarity between subgenomes, introgressions from a dominant genome, and other issues complicate accurate mapping and variant calling. The recent publication of a full chromosome length reference genome and improvements of deep sequencing for polyploids facilitated overcoming part of those difficulties, including the development of linkage maps. A biparental population was generated crossing the NCSU selections NCS 10-080 × NCS 10-147. The parents and 280 seedlings were sequenced using the reduced representation sequencing OmeSeq protocol resulting in 2.47 billion reads. The ngsComposer application was used for quality control, demultiplexing, and filtering, resulting in 1.84 billion reads. The read alignment resulted in a coverage of 92.32% of the four subgenomes in the allo-octoploid strawberry reference genome 'Camarosa' 1.0. The map construction was done using the MAPpoly R package. After a quality control screening, a total of 6133 markers and 212 offspring individuals were used to build a genetic linkage map comprised of 28 linkage groups with a total length of 3154 cM and 4022 SNP makers. The minimum linkage group size was 30.2 cM and 49 markers, and a maximum of 155.69 cM and 168 markers, with an average interval genetic distance of 1.32 cm. The high marker density and the correspondence between the number of assembled linkage groups and the number of expected chromosomes indicates that MAPpoly R is a robust analysis. This analysis provides an excellent framework map for forthcoming studies, including QTL analysis and understanding modes of inheritance in this complex polyploid species.

Breakout room #6

Genomic prediction for yield and processing traits in the tetraploid potato

Jeewan Pandey[1], *Douglas C. Scheuring*[1], *Jeffrey W. Koym*[2], and *Maria Isabel Vales*[1]. [1]Department of Horticultural Sciences, Texas A&M University, College Station, TX. [2]Texas A&M AgriLife Research and Extension Center, Lubbock, TX.

The current breeding process to develop a new potato cultivar takes a long time (10–15 years). One way to speed up the process and make it more efficient is to shorten the recurrent selection breeding cycle. This can be achieved by assigning breeding values to clones in the breeding pipeline and bringing those with the most favorable breeding values as parents for the crossing block. The aim of this study was to implement one angle of genomic selection by obtaining breeding values of chipping potato clones and recommend parents for the breeding program. Five hundred and forty-nine unique chipping potato clones were evaluated between 2017 and 2020 near Dalhart, TX, and genotyped using the Illumina Infinium Potato SNP array. Genomic-estimated breeding values (GEBVs) for chip color, chip quality, specific gravity, and total yield were obtained using the package StageWise. Potato clones with the most favorable GEBVs were identified and recommended as parents. The mean reliability of the GEBVs

obtained were 0.75, 0.43, 0.61, and 0.33 for chip color, chip quality, specific gravity, and total yield, respectively. Breeders will increase the probability of transferring useful traits from parents to their progeny by choosing parental lines with the most favorable GEBVs. In turn, progeny with the best GEBVs can be re-used as parents or advanced to become new varieties. Thanks to the development of software packages suitable for polyploid species, genomic selection in potatoes is becoming more feasible and attractive.

Breakout room #7

Exploring the Genetic Control of Potato Tuber Dormancy

Ao Jiao[1], Sanjeev Gautam[1], Jeewan Pandey[1], Douglas C. Scheuring[2], Jeffrey Koym[1], and M. Isabel Vales[1]. [1]Department of Horticultural Sciences, Texas A&M University, College Station, TX. [2]Texas A&M AgriLife Research and Extension Center, Lubbock, TX.

Potato tuber dormancy is defined as the period after harvest during which tubers do not sprout even under favorable conditions. The length of the dormancy period is measured from vine kill until tubers start sprouting and it is affected by genotype and environmental factors. Premature dormancy break is a major factor causing post-harvest tuber quality reduction. Common methods to prevent sprouting include cold storage and the use of sprout inhibitors. Cold storage causes cold-induced sweetening and results in higher energy costs, whereas sprout inhibitors raise health and environmental concerns. Developing potato varieties with long dormancy could contribute to reducing the use of cold storage and sprout inhibitors. In this study, we evaluated tuber dormancy variation and investigated the genetic background of tuber dormancy. Over 200 clones from the Texas A&M Potato Breeding Program were grown in Dalhart, TX in 2019 and were evaluated for dormancy at room conditions (18 °C, RH 60%, dark). The clones exhibited variation in dormancy ranging from 38 to 155 days, with the Russets having significantly longer dormancy (> 96 days) than other market groups (70 – 80 days). Two Texas A&M varieties, Reveille Russet and Vanguard Russet, were among the clones with the longest dormancy. A genome-wide association study was performed using GWASpoly with Infinium Illumina 22K V3 Potato Array to identify genomic regions associated with tuber dormancy. The main QTL identified was on chromosome 9, explaining 11% of the phenotypic variation. Follow-up evaluations will be conducted at additional locations under room temperature and cold storage.

Breakout room #8

Development of Genomic Resources for Cultivated Blueberry

Ira A. Herniter and Nicholi Vorsa. Rutgers University, New Brunswick, NJ.

Blueberry (Vaccinium sect Cyanoccocus) is an increasingly important fruit crop native to North America. Collected from the wild for thousands of years, blueberry was only domesticated in the early twentieth century. While blueberry has long been a minor crop, recently interest in its health properties, including high levels of anthocyanins, has led to increased consumption and cultivation around the globe. However, quality genetic resources for blueberry have been lacking, greatly hampering the development of new varieties which can produce well under the constraints of changing climatic and pest regimes. We report the development of several populations being used for trait mapping. Four are diploid interspecific populations resulting from crosses between V. corymbosum, native to temperate climate, and V. darrowi, native to a subtropical climate, as well as a large germplasm collection consisting of released blueberry cultivars and wild relatives, including species from across North America, as well as species from Europe, Pacific Russia, and South America. The populations have great diversity in a range of traits, including leaf shape, berry size, fruit chemistry, flowering time, among others. The populations are an important resource for the development of new varieties and increasing understanding of blueberry physiology.

Breakout room #9

Improving zoysiagrass (*Zoysia Willd. species)* for drought tolerance across the Southern US

Beatriz Tome Gouveia[1], Brian M. Schwartz[2], Yangi Wu[3], Kevin E. Kenworthy[4], Ambika Chandra[5], Paul L. Raymer[2], Marta T. Pudzianowska[6], Esdras M. Carbajal[1], Manual R. Chavarria Sanchez[4], Jing Zhang[2], Bradley M. Batterhsell[3],

Pamela S. Rowe, Meghyn B. Meeks[5], Tianyi Wang[5], Chase N. McKeithen[4], and Susan R. Milla-Lewis[1]. [1]North Carolina State University, Raleigh, NC. [2]University of Georgia, Tifton, GA. [3]Oklahoma State University, Stillwater, OK. [4]University of Florida, Gainesville, FL. [5]Texas A&M University, Dallas, TX. [6]University of California, Riverside, CA.

Zoysiagrass (Zoysia Willd. species, 2n = 4x = 40), a complex that encompasses 11 species, are primarily used for home lawns, public parks, and athletic fields, being the second most used warm-season turfgrass on golf courses in the US. Water limitations are currently one of the biggest challenges for the turfgrass industry. Starting in 2010, a Turfgrass Specialty Crop Research Initiative (SCRI) project, funded by USDA-NIFA, has focused on addressing the problems of limited availability and reduced quality of water for irrigating turfgrass areas by breeding warm-season turfgrass species for improved drought and salinity tolerances. The objective of this study was to evaluate the performance of zoysiagrass breeding lines from the breeding programs at University of Georgia – Tifton and Griffin, North Carolina State University, Texas A&M University and University of Florida under drought conditions. Field trials arranged as randomized complete-block designs with three replications were installed at research facilities at Citra, FL, and Stillwater, OK in the summer of 2020. The response variables evaluated were turfgrass quality under normal or non-drought conditions (TQND), and both percent green cover (PGC), evaluated using UAS, and turfgrass quality (NTEP ratings from 1 to 9) (TQD) under drought conditions. The genetic variance was significant and non-significant for all traits in the single-environment and multi-environment analysis, respectively. The genotype-by-environment interaction variance was significant for all traits. Heritability estimates were above 0.50 for all traits in all locations, except for TQD in Citra. A high positive correlation was observed between TQD and PGC in both locations, whereas these traits showed low correlation with TQND. Several breeding lines performed better than the checks for both TQD and PGC at both locations. Evaluation of these genotypes will continue through 2023. Results of this study will support the selection of drought-tolerant elite zoysiagrass genotypes with increased performance stability for the target regions.

Breakout room #10

Genome-Wide Association Study on Potato Tuber Defects Under Heat Stress

Sanjeev Gautam and M. Isabel Vales. Department of Horticultural Sciences, Texas A&M University, College Station, TX. Texas A&M AgriLife Research and Extension Center, Lubbock, TX.

Heat stress reduces marketable tuber yield and quality of potatoes. Tuber defects can be external (heat sprouts, chained tubers, knobs) or internal (vascular discoloration, internal heat necrosis). Successful cultivation of potatoes under heat stress requires planting heat-tolerant varieties that can produce high yields of marketable tubers. Heat tolerance is a complex trait and breeding for it possesses several bottlenecks. To facilitate future marker-assisted selection for heat tolerance, a genome-wide association study (GWAS) aimed to identify genomic regions associated with heat tolerance was conducted. Phenotyping for a panel of 217 diverse potato genotypes was conducted near Springlake, Texas (heat stress location) for two years using a randomized complete block design with two replicates. The genotypes differed in their capacity of expressing the external as well as internal defects on tubers under heat stress. GWAS was conducted using GWASpoly with Infinium Illumina 22 K V3 Potato Array. Significantly associated SNPs with external defect traits were located on chromosomes 3, 4, 6, and 7 while those with internal defect traits were located on chromosomes 3 and 10. The identified genomic regions may be important to improve heat tolerance in potatoes. Fine mapping of identified regions and validation of the markers associated with these regions would be required to further understand the mechanism involved in heat tolerance.

Breakout room #11

Sequenced Genomes of Allotetraploid *Poa annua* and its Diploid Progenitors, *Poa infirma* and *Poa supina*, Provide Insight into the Evolution and Breeding of Versatile Polyploid

Christopher Benson[1], David Huff[1], Shaun Bushman[2], Peter Maughan[3], Rick Jeelen[3], and Matthew Robbins[2]. [1]Pennsylvania State University, State College, PA. [2]USDA-ARS, Logan, UT. [3]Brigham Young University, Provo, UT.

Poa annua (annual bluegrass, 2n=28, AABB) is an allotetraploid grass species that outperforms its diploid progenitors in both diversity of morphologies and geographic range. On golf course putting greens, Poa annua's ability to produce seed under 3mm mowing height has contributed to its discordant reputations as both a noxious weed and a valued commodity. Despite an estimated $40 billion U.S. turfgrass industry, there have been limited successful efforts directed at managing Poa annua, either for or against, primarily due to its complex genetic and

epigenetic versatility. Here we present the pseudomolecule-level genome assemblies of Poa annua and its diploid progenitors, Poa infirma, and Poa supina. BRAKER2 annotations of these species with Iso-Seq RNA-evidence yielded 72,034, 37,207, and 35,698 high-confidence proteins, respectively. Both the assemblies and the annotations for all species contained >90% conserved orthologs, corroborating their quality. We demonstrate that the parental diploid genomes accurately represent the A and B subgenomes of Poa annua and characterize genetic exchange between and within the subgenomes of Poa annua. We show that the bifurcating ecological niches of the parents is mirrored by genomic and structural mutations in their diploid genomes. The subgenomes of Poa annua bear strong resemblance to its progenitors, confirming its status as a neo-allotetraploid. We speculate that Poa annua's global proliferation is conferred through the union of two parental genomes with wide genetic distance for hybridization and contrasting ecological ranges. We plan to incorporate genomic and transcriptomic resources to aid in better targeting of Poa annua in turfgrass beeding applications.

Breakout room #12

Development of a Genotyping by Sequencing Pipeline in Tetraploid Roses (*Rosa sp.*)

Tessa Hochhaus, Cristiane H. Taniguti, Jeekin Lau, Patricia E. Klein, David H. Byrne, and Osar Riera-Lizarazu. Department of Horticultural Sciences, Texas A&M University, College Station, TX.

Roses are highly heterozygous and most commonly diploids, triploids, and tetraploids. Genotyping by sequencing (GBS) has been performed in diploid rose populations, however, it has not been done in populations with higher ploidy because of their increased complexity (autopolyploidy). This complexity is due to the greater number of genotypic classes and the difficulty in accurately calling allele dosage. GBS uses restriction enzymes to reduce genome complexity and adapter barcodes to allow the pooling of multiple samples to increase the efficiency and to lower the sample cost. In this study, we are optimizing a GBS protocol for tetraploid roses using three populations (Morden Blush x George Vancouver, Stormy Weather x Brite Eyes, and Brite Eyes x My Girl). The optimization will entail varying sequencing read depth and coverage, while minimizing missing data, and using in-house workflows to test various combinations of open-source software for quality control, alignment of reads, identifying SNPs, and dosage calling. Through the development of this pipeline we hope to facilitate cost-effective genotyping in polyploid roses and the use of genomic-assisted breeding.

Breakout room #13

Identifying a Rose Germplasm Panel to Attain Optimal SNP Array Genotype Calling of Small Samples of Genotyped Individuals

Jeekin Lau, Cristiane H. Taniguti, David Byrne, and Oscar Riera-Lizarazu. Texas A&M University, College Station, TX.

Since genotyping with the Axiom WagRhSNP68K SNP array can be cost prohibitive, we explored an approach that would permit robust genotyping of samples in one or two 96-well plates. We have observed that genotyping accuracy via SNP arrays increases as the number of individuals used for genotype calling increases. We reasoned that this increased accuracy may be due to greater sample size and allelic diversity. To test this idea, we conducted an experiment where one bi-parental mapping population of 94 individuals plus two parents were clustered alone (one plate of genotyping) and in combinations with sets of related biparental populations and unrelated germplasm with increasing numbers and various levels of genetic diversity. We then compared both marker statistics and the linkage map quality generated from genotype calls of the target mapping population using the various datasets. As the number of individuals used in clustering increased, the number of useful markers increased nominally. However, the resulting linkage maps revealed that the addition of other genotypes in the marker clustering step resulted in shorter total map length and smaller gap sizes as the number of individuals and diversity increased. The decreased map lengths and gap sizes indicate that the inclusion of other genotypes helped genotyping accuracy. The output of this study will be a core set of genotyped rose germplasm that may be used to improve genotype calling of small samples of genotyped materials.

Breakout room #14

Unusual dalliances within the polyploid Chrysanthemum species complex

Neil O. Anderson[1], Liesl Bower-Jernigan[1], Rajmund Eperjesi[1], Robert Suryani[1], Steven Gullickson[2], and Albert Radloff[2]. [1]University of Minnesota, Minneapolis, MN. [2]MGK, Minneapolis, MN.

Many species in the Chrysanthemum complex and its alliance of satellite genera (e.g. Arctanthemum, Leucanthemum, Tanacetum) have been bred and selected since the 15th century BCE. Crops include a diversity of uses and forms from green pesticides (pyrethrum, C. cinerariifolium, C. coccineum; 2n=2x=16, 2n=3x=24), edible shoots (C. carinatum, C. coronarium, C. segetum; 2n=2x=16), salt tolerance (C. arcticum, subsp. Arcticum, polaré; 2n=2x=16, 2n=4x=32), to ornamentals (cut flower, potted plant, garden types, C. xgrandiflorum, C. xhybridum; 2n=6x=54) also used for medicinal, herbal teas, and wine. The most widely cultivated crop, C. xgrandiflorum, is a complex perennial geophyte, an allohexaploid of >10 species (C. zawadskii, etc.), complicated by self incompatibility (3 S loci), pseudo-self compatibility (PSC), aneuploidy, sterility, inbreeding depression, and genetic load. Chrysanthemum populations in the University of Minnesota breeding program were analyzed for genetic structure within/among species and cultivar series to determine alliances within the ploidy complex. Genotypic analyses (GBS; DarTseqLD) was used to identify 389 low density, unique SNPs and determine genetic structure within and among wild (C. arcticum, subsp. Arcticum, polaré; C. zawadskii) and cultivated (C. cinerariifolium, C. xgrandiflorum 'Minn' series, C. xhybridum 'MammothTM' series) species populations. Principal Coordinates Analysis (PcoA) of all species showed two clusters of C. arcticum/C. cinerariifolium and C. xgrandiflorum/C. xhybridum/C. zawadskii with 74.9% diversity for the principal component (PcoA1) and 8.1% for PcoA2. Surprisingly, the first subgroup had C. cinerariifolium in close dalliance with C.a. subsp. Polaré (Nome, AK), followed by C.a. subsp. Arcticum (Aleutian Islands), C. arcticum (Anchor Point, Kenai, Ninilchik, Valdez, AK) with C.a. subsp. Polaré (Churchill, Manitoba, Canada). Chrysanthemum xgrandiflorum had the greatest level of genetic diversity, although it slightly overlapped with both C. zawadskii and C. xhybridum. The 'Minn' and 'MammothTM' series had low levels of genetic differentiation due to C. xhybridum being derived from C. xgrandiflorum and C. weyrichii (2n=6x=54). Future research will focus on phenotypic trait/SNP associations in a genome-wide association study (GWAS) to aid in marker-assisted selection.

Breakout room #15

Quantitative Trait Loci Associated with Flower Color Transition Phenotype in Tetraploid Roses

Haramrit Gill, Jeekin Lau, Qiuyi Fu, Natalie Anderson, David H. Byrne, and Oscar Riera-Lizarazu. Texas A&M University, College Station, TX.

Flower color is one of the most important breeding traits in ornamental roses. The combination of particular anthocyanins, their co-factors and their concentrations leads to different pigmentation patterns. We observed an interesting characteristic that we call 'flower color transition' in two tetraploid rose populations a) 'Stormy Weather' (SW) X 'Brite Eyes' (BE), b) 'My Girl' (MG) X 'Brite Eyes' (BE). The roses that exhibit this phenotype have flowers that transition from a light-yellow color to a dark pink/red (accumulation of anthocyanins) as the flower ages leading to bushes peppered with flowers of multiple colors. To our knowledge, the genetic control of this phenotype has not been studied in roses previously. Here, we present studies to better understand the inheritance of this trait and to identify quantitative trait loci (QTL) in two tetraploid bi-parental populations segregating for the 'flower color transition' trait. Our analysis suggests the presence of QTL on chromosomes 3 and 4. The location of QTL identified for the flower color transition coincides with the location of some genes involved in flavonoid biosynthetic pathway. Additional studies are underway to validate these results.

Breakout room #16

Predicting Length to Width Ratio, Roundness, and Compactness in Potatoes

Michael Miller and Laura Shannon. University of Minnesota, Saint Paul, MN.

Potato is an important crop to the global food system; however, adoption of new potato varieties has been slow compared to many other staple food crops. This is in part due to the importance of many traits beyond yield, in particular quality traits, to the marketability of potatoes. Quality traits are often measured using subjective and imprecise visual scales. These scales introduce error due to rater fatigue, rater experience, and differences in scale interpretation between raters. Visual scales also limit differentiation between potato clones which express a trait at similar but not identical levels. We have developed an image analysis platform in the R programming language to provide objective and precise numeric measurements of several potato tuber quality traits. Among these traits are measurements of tuber shape including length to width ratio, roundness, and compactness. Combining the phenotype data provided by this platform with the genomic selection tools provided by Tools for Polyploids could

allow for robust genomic selection models of potato tuber shape quality traits. A collection of 82 chip market class potato clones from the University of Minnesota breeding program were evaluated for shape traits over 3 field seasons. Genotyping was preformed using the SolCAP SNP array. Allele dosage calling was performed using the fitPoly R package. The StageWise package was then used to create predictions of genomic estimated breeding values for tuber shape traits.

Breakout room #17

Multiploidy support in polyRAD

Lindsay V. Clark, Joyce Njuguna, Alexander E. Lipka, and *Erik J. Sacks.* Department of Crop Sciences, University of Illinoism, Urbana-Champaign, Urbana, IL.

polyRAD is an R package for Bayesian genotype calling from sequence read depth in diploid and polyploid organisms. It can use population structure or mapping population design to inform genotype calls, and can export discrete or continuous genotypes. Although the original version of polyRAD allowed inheritance mode to vary across the genome, it still required all individuals to be the same ploidy, limiting its use in staple crops such as banana and yam in which breeding populations typically consist of a mixture of ploidies. polyRAD 2.0 will support multiploidy, allowing simultaneous genotyping of individuals of different ploidies. The "possiblePloidies" slot will still be used to indicate potential inheritance modes for loci. A new slot called "taxaPloidy" contains one integer for each individual to indicate its ploidy, and acts as a multiplier for the values stored in "possiblePloidies". Examples of how to code this information in various crops will be presented in the digital poster. We will also present Miscanthus sacchariflorus as a use case, in which introgression has occurred among diploid, triploid, and tetraploid populations. The development version of polyRAD 2.0 can be installed from GitHub.

Breakout room #18

Increasing the Prediction Accuracy in Genomic Selection of Complex Traits using WGBLUP

Cesar A. Medina[1], Harpreet Kaur[2], Ian Ray[2], and Long-Xi Yu[1].
[1]United States Department of Agriculture-Agricultural Research Service, Plant Germplasm, Introduction and Testing Research, Prosser, WA. [2]Department of Plant and Environmental Sciences, New Mexico State University, Las Cruces, NM.

Genomic selection (GS) is a variant of marker-assisted selection, in which genome-wide markers are used to determine the genomic estimated breeding value (GEBV) of individuals in a population to a specific trait. GS is useful in complex traits controlled by many genes with small effects. However, some complex traits such as biomass yield or abiotic stress tolerance have low prediction accuracy (measured as Pearson correlation between GEBV and phenotypic values). There is a need to increase the prediction accuracy to employ GS in breeding programs. In this work, we developed and tested an alternative GS model named weighted GBLUP (WGBLUP). We integrated DNA marker significant values of genome-wide association studies (GWAS) in WGBLUP analyses. We performed a case study using phenotypic data on biomass yield under salt stress of alfalfa and 13 phenotypic traits of potato to validate the WGBLUP model. This approach increased prediction accuracies from 50% to more than 80% for alfalfa yield under salt stress and up to 90% in potato tuber length. The use of the WGBLUP model will allow to implement GS in different breeding programs, increasing the selection accuracy in complex traits.

Breakout room #19

A Semi-Automated SNP-Based Approach for Contaminant Identification in Biparental Polyploid Populations of Tropical Forage Grasses

Felipe Bitencourt Martins[1], Aline Costa Lima Moraes[1], Alexandre Hild Aono[1], Rebecca Caroline Ulbricht Ferreira[1], Lucimara Chiari[2], Rosangela Maria Simeão[2], Sanzio Carvalho Lima Barrios[2], Mateus Figueiredo Santos[2], Liana Jank[2], Cacilda Borges do Valle[2], Bianca Baccili Zanotto Vigna[2], Anete Pereira de Souza[1]. [1]University of Campinas, Sao Paulo, Brazil. [2]Embrapa Gado de Corte, Mato Grosso do Sul, Brazil.

Artificial hybridization plays a fundamental role in plant breeding programs since it generates new genotypic combinations that can result in desirable phenotypes. Depending on the species and mode of reproduction, controlled crosses may be challenging, and contaminating individuals can be introduced accidentally.

In this context, the identification of such contaminants is important to avoid compromising further selection cycles, as well as genetic and genomic studies. The main objective of this work was to propose an automated multivariate methodology for the detection and classification of putative contaminants, including apomictic clones, self-fertilized individuals, half-siblings and full contaminants, in biparental polyploid progenies of tropical forage grasses. We established a pipeline to identify contaminants in genotyping-by-sequencing (GBS) data encoded as allele dosages of single nucleotide polymorphism (SNP) markers by integrating principal component analysis (PCA), genotypic analysis (GA) measures based on Mendelian segregation and clustering analysis (CA). The combination of these methods allowed the correct identification of all contaminants in all simulated progenies (n=200) with more than 690 markers and the detection of putative contaminants in three real progenies of tropical forage grasses, providing an easy and promising methodology for the identification of contaminants in biparental progenies of tetraploid and hexaploid forages or other species. The proposed pipeline was made available through the polyCID Shiny app, which was developed in R language with a user-friendly interface designed to facilitate its use by plant breeders. Furthermore, it can be easily coupled with traditional genetic approaches, such as linkage map construction and other SNP based techniques, thereby increasing the efficiency of breeding programs.

Sponsor Acknowledgments

Neuhouse Farms

Collaborator Acknowledgments

Neuhouse Farms

Woolf Roses L.L.C.