nucleotide fasta sequences, multiple sequence alignments, and inferred maximum likelihood phylogenies. Optionally, a user can provide externally predicted coding sequences and their corresponding amino acid translations derived from a transcriptome assembly or gene predictions from a sequenced genome. PlantTribes is freely available at https://github.com/dePamphilis/PlantTribes.

P0346: Bioinformatics: Software

Highly Interactive Data Analysis and Productivity (HIDAP) tools for Breeding

Reinhard Simon[1], Omar Benites[1], Raul Eyzaguirre[1], Awais Khan[2] and Luka Wanjohi[1], (1)International Potato Center (CIP), Lima, Peru, (2)International Potato Center, Lima, Peru

Plant breeders routinely work with large datasets that increasingly also add new levels of dimensionality and complexity using, for example, –omics data. This gives breeders additional work for data analysis, which can be addressed at least partially using computational tools. To some extent, field trials are routine and repetitive and can benefit from standardized reporting; high density data sets and complex relationships can benefit from visualization and visual navigation; and complex relationships can benefit from applying linked data concepts and interactivity. Here we apply a new generation of R based libraries to implement a custom platform (HIDAP) for productively handling such breeding data. The HIDAP platform consists of a set of R libraries addressing tasks like data quality control, basic analysis of single and multiple breeding trials, and integrating environmental, phenotypic and genomic data. It also creates printable reports based on modifiable and extensible templates. It can be used as a companion tool with software platforms like the Integrated Breeding Platform (IBP) and the Sweetpotatobase since it can directly exchange files with these two and exchange data in general *via* implementing the 'Plant Breeding Application Programming Interface API'. Like R, the platform is fully open source and available across operating systems, and can be used offline and online. The libraries are available from a dedicated mini CRAN server, and can also be used individually.

P0347: Bioinformatics: Software

GACD: Integrated Software for Genetic Analysis in Clonal F1 and Double Cross Populations

Luyan Zhang[1], Jiankang Wang[2], Lei Meng[2] and Wencheng Wu[2], (1)Instititue of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China, (2)Chinese Academy of Agricultural Sciences, Beijing, China

Clonal species are common among plants. Clonal F_1 progenies are derived from the hybridization between two heterozygous clones. In self- and cross-pollinated species, double crosses can be made from four inbred lines. A clonal F_1 population can be viewed as a double cross population when the linkage phase is determined. The software package GACD (Genetic Analysis of Clonal F_1 and Double cross) is freely-available public software, capable of building high-density linkage maps and mapping quantitative trait loci (QTL) in clonal F_1 and double cross populations. Three functionalities are integrated in GACD version 1.0: binning of redundant markers (BIN); linkage map construction (CDM); and QTL mapping (CDQ). Output of BIN can be directly used as input of CDM. After adding the phenotypic data, the output of CDM can be used as input of CDQ. Thus, GACD acts as a pipeline for genetic analysis. GACD and example datasets are freely available from www.isbreeding.net.

P0348: Bioinformatics: Software

Field Book: An Open-Source Android App for Collecting Phenotypic Data

Trevor W. Rife and Jesse Poland, Kansas State University, Manhattan, KS

Plant breeding and genetics research is an inherently data-driven enterprise. Typical experiments and breeding nurseries contain thousands of unique entries and programs will often evaluate tens of thousands of plots each year. To operate efficiently on this scale, electronic data management becomes essential. Many research programs, however, continue to operate by scribing and transcribing massive amounts of data on paper field books. This form of data management places heavy burdens on human resources, decreases data integrity, and limits future utilization of data. To help address these constraints, we have developed Field Book, an open-source app for electronic data capture that runs on consumer-grade Android phones and tablets. By focusing on a simple, stand-alone application with an intuitive interface, we attempt to decrease both technological and cost barriers that hinder adoption of electronic data management in breeding programs. The simplicity of Field Book allows adoption without a steep learning curve. With low-cost, accessible solutions, the vision of one handheld per breeder can become a reality for breeding programs around the world. Transformational capacity in electronic data collection and management will be essential to realize a contemporary green revolution.

P0349: Bioinformatics: Software

TheSNPpit - A Multi Panel SNP Data Management Package

Eildert Groeneveld, Institute of Farm Animal Genetics (FLI), Neustadt, Germany and Helmut Lichtenberg, Institute of Farm Animal Genetics (FLI), Neustadt am Rbge, Germany

TheSNPpit is a database system for managing multi panel SNP genotype data. It has implemented three new ideas: highly compressed vector storage in a database, set based manipulation, and a very fast export written in C with Perl as the base for the framework and PostgreSQL as the database backend. Its novel subset system allows the creation of named subsets based on the definition of SNP (through MAF, no calls, and chromosomes) and sample lists at basically no storage costs, thus avoiding the issue of proliferating file copies. The named subsets are exported for down stream analysis.

Plink ped and map files are processed as inputs and outputs with a 2 bit storage of SNPs in the database allowing to fill 4 mio SNPs/1MB. TheSNPpit can handle SNP data from panels of any size. To investigate performance scaling, a database with more than 18.5 mio samples has been created with 3.4 trillion SNPs from 12 panels ranging from 1000 through 20 mio SNPs resulting in a database of 850GB. The performance scales linearly with the number of SNPs and is largely independent from panel and database size. Import speed is around 6 mio SNPs/sec, export between 60 and 120 mio SNPs/sec.

A completely generalized procedure allows storage of phenotypes, which will automatically be exported with the SNP genotypes for further downstream processing. Being command line based, imports and exports can easily be integrated into pipelines. TheSNPpit is available under the Open Source model.