UNIVERSITY OF GEORGIA
College of Agricultural & Environmental Sciences

Chavez Lab

# Target genotyping using Capture-Seq technology to characterize the population structure and phylogenetic relationships of peaches from Australia and other parts of the world.

Caitlin McCann[1], Jose X. Chaparro[2], Bruce Topp[5], Dougal M. Russell[6], Jeekin Lau[3], Jordan Knapp-Wilson[1], Rachel Itle[4], and Dario Chavez[1,4]

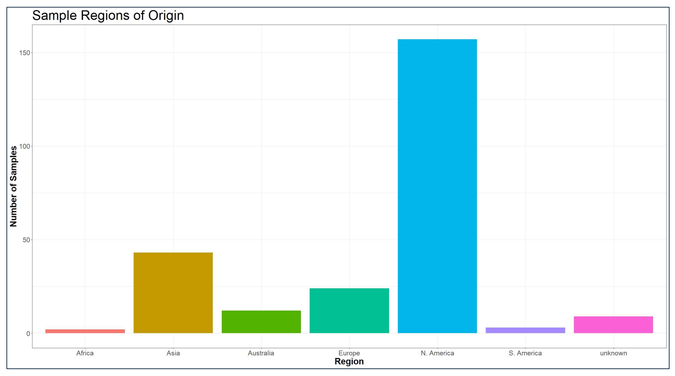
[1]Institute of Plant Breeding Genetics, and Genomics. University of Georgia, Griffin, GA 30223 [2]Horticultural Sciences Department. University of Florida, Gainesville, FL 32611.
[3]Southeastern Fruit and Tree Nut Research Laboratory. USDA-ARS, Byron, GA 31008. [4]Department of Horticulture. University of Georgia, Griffin, GA 30223.
[5]Cemtre for Horticultural Science, University of Queensland, QLD 4067, Australia. [6]Queensland Government, Department of Agriculture and Fisheries, QLD 4000, Australia

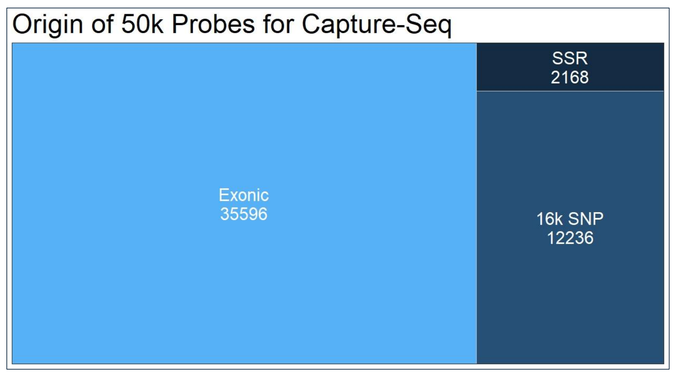
## Background

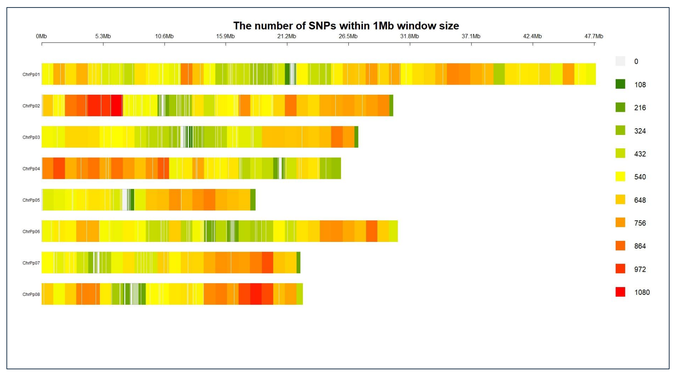
**Germplasm Diversity**
- Peaches in the US are highly homozygous
  - A genetic bottleneck in the 1870s resulted in low diversity in US peaches
  - Acquiring new diversity would bolster breeding efforts and help combat disease and climate related challenges
  - Therefore, identifying potential sources of germplasm diversity is a priority

**SNP Markers**
- The current SNP chip array from Illumina is inflexible and expensive
  - One must buy the whole array, even if you only need a few SNPs
  - A probe panel would allow researchers to choose their targets and would expand on the SNP markers currently available
  - Genotyping SNP-based methods such as KASP can be done in-house

**Our Goals**:
- Goal 1: Determine whether Australian peach populations are related to Chinese populations
- Goal 2: Develop regionally-informative SNP markers based on 50k probe panel for genotyping



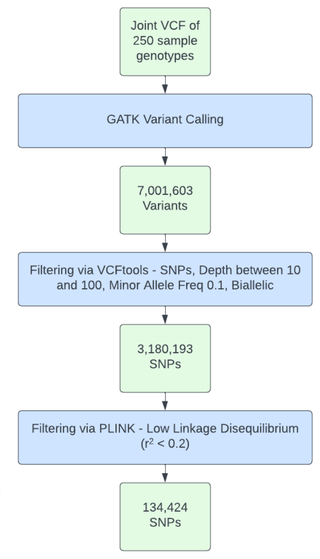



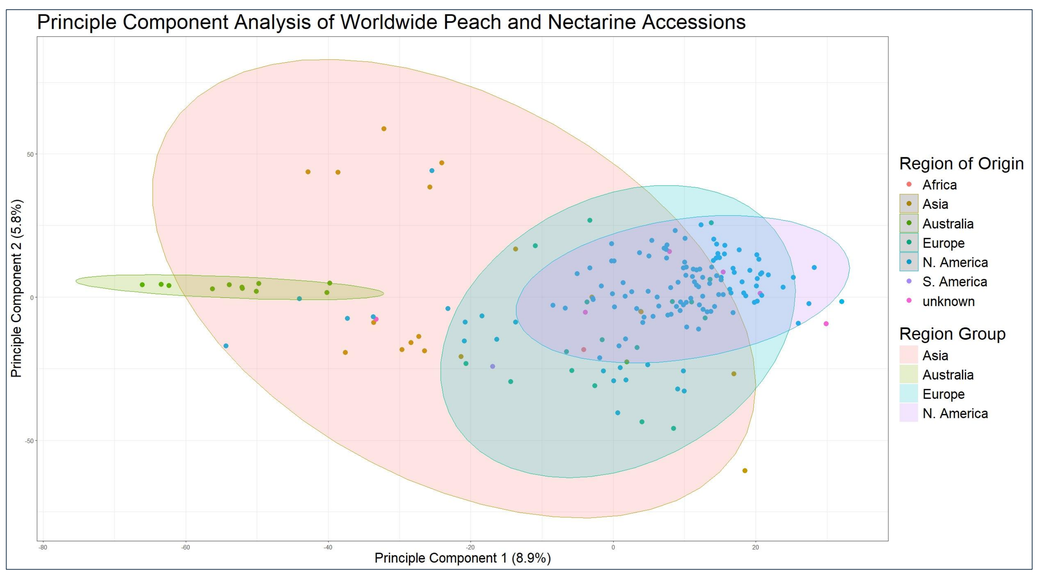

## Methodology

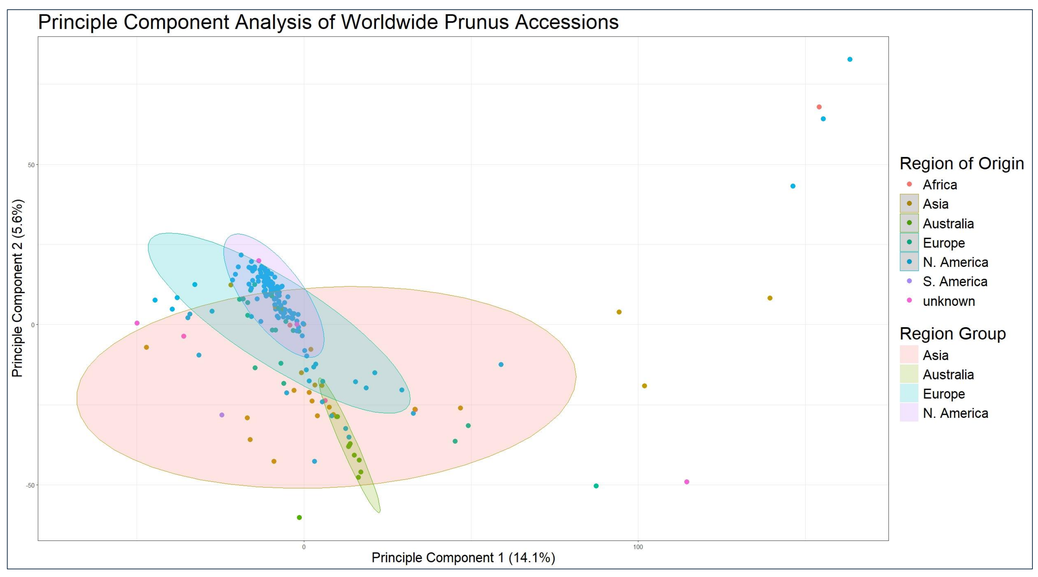
**Sample Acquisition**
- Dempsey Farm, Griffin, GA
- USDA-ARS GRIN repository
- Australian Collaborators

**Genotyping**
- Capture-Seq through LGC BioSearch Technologies
- 50k probes developed based on the existing 16k SNP array, SSR markers, and exonic regions

**Data Analysis**
1) Evaluate data quality
   - FastQC and MultiQC
2) Align reads to genome and join all into one .vcf
   - VCFtools
3) Call variants (indels and SNPs)
   - GATK
4) Filter variants
   - SNPs, read depth between 10 and 100, MAF 0.1, low LD, biallelic
   - VCFtools, PLINK
5) Analyze results
   - PCA, Structure, heterozygosity, missing data



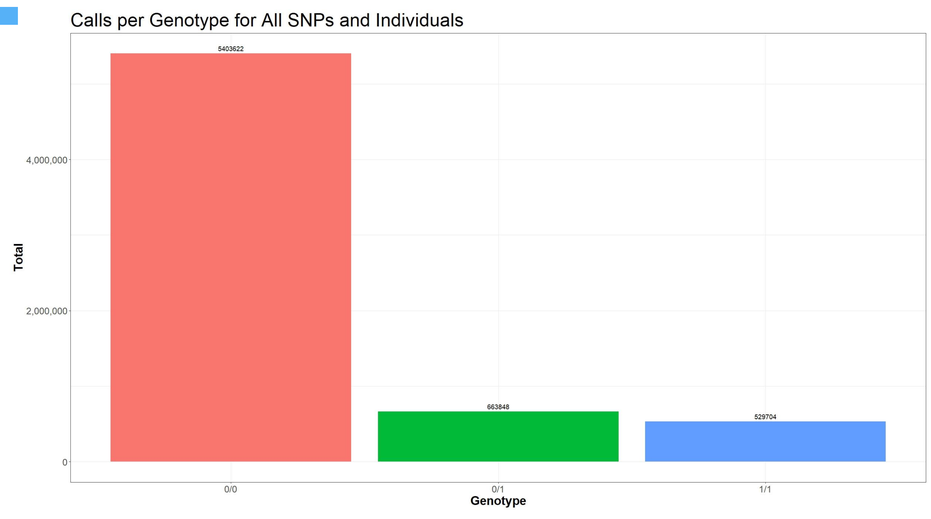




- **221 - *P. persica***
  - **196 Peaches**
  - **25 Nectarines**
- **29 - other species**
  - ***P. augustifolia***
  - ***P. avium***
  - ***P. davidiana***
  - ***P. dulcis***
  - ***P. mira***
  - ***P. salicina***
  - ***P. domestica***
  - ***P. kansuensis***
  - ***P. mume***
  - **hybrids**

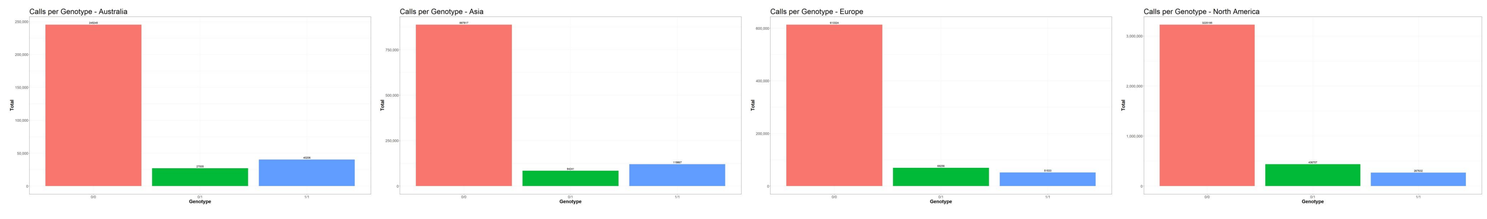
## Results

- 134,424 SNPs after filtering
  - ~11k SNPs overlap with 16k SNP array
  - All SSR-based probes yielded SNPs
- Clear separation of regional accessions
  - Next Step: isolate regionally informative SNP markers
  - Next Step: use these markers to further evaluate Australian germplasm
- 48 individuals with high percent missing data (>60%)
  - To be re-sequenced




Chavez Lab

Personal Website




## Acknowledgements

We would like to acknowledge our collaborators in Queensland, Australia for supplying us with samples, as well as the USDA-ARS GRIN repository.
This study was supported in part by resources and technical expertise from the Georgia Advanced Computing Resource Center, a partnership between the University of Georgia's Office of the Vice President for Research and Office of the Vice President for Information Technology.