RosBREED

Enabling marker-assisted breeding in Rosaceae

www.rosbreed.org

A Coordinated Approach to Peach SNP Discovery in RosBREED

Nahla Bassil[1] - Presentor

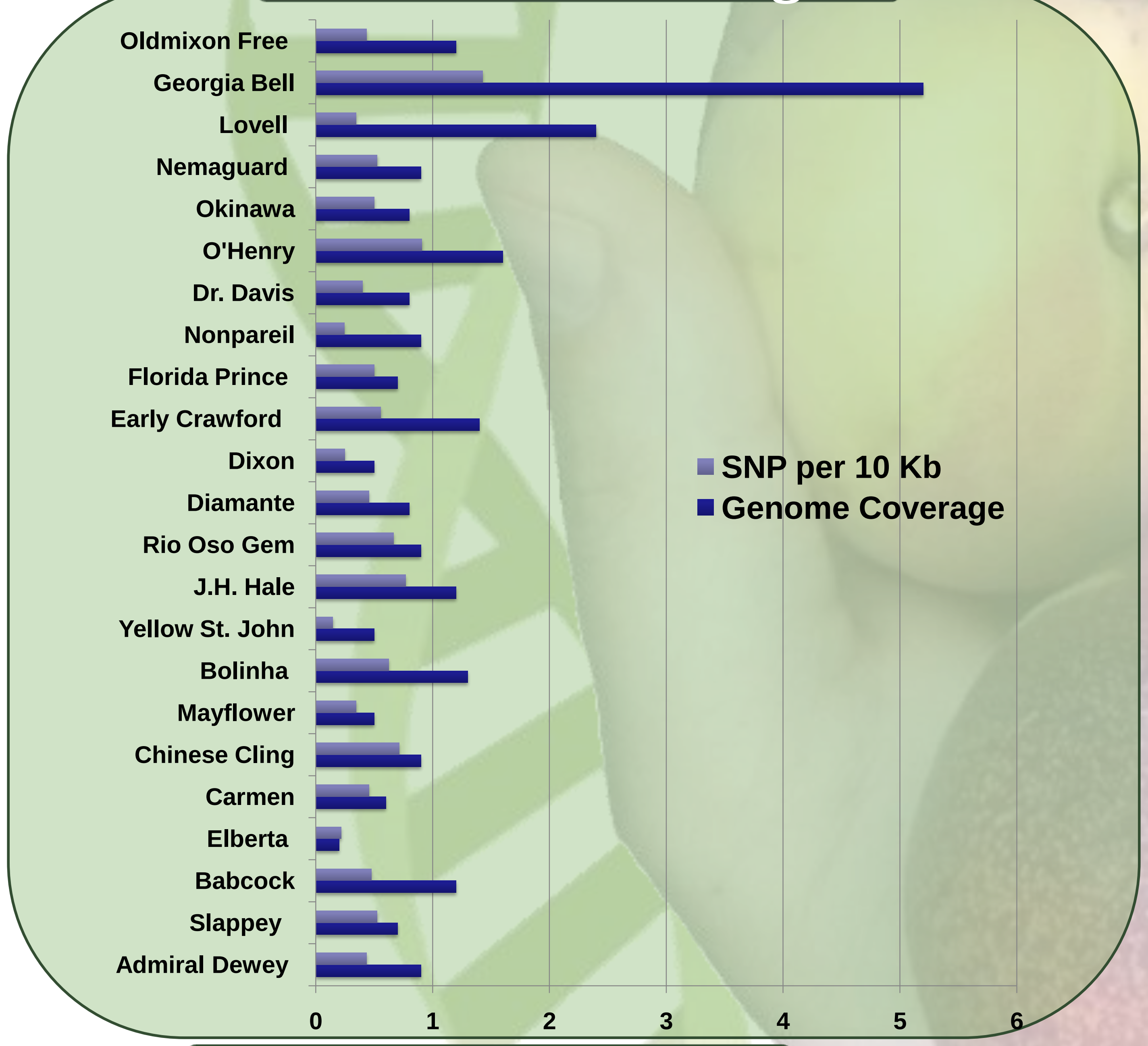
Barbara Gilmore[1], Ignazio Verde[2], Bryon Sosinski[3], Pere Arus[4], Gennaro Fazio[5], Ksenija Gasic[6], John Clark[7], David Byrne[8], Tom Gradziel[9], Dorrie Main[10], Michele Morgante[11], Larry Wilhelm[12], Todd Mockler[12], Cameron Peace[10], Amy Iezzoni[13]

[1]USDA-ARS, National Clonal Germplasm Repository, Corvallis, OR 97333 USA
[2]CRA - Centro di Ricerca per la Frutticoltura, 00134 Roma, Italy
[3]Department of Horticultural Science, North Carolina State University, Raleigh, NC 27695, USA
[4]Institut de Recerca i Tecnologia Agroalimentàries, , 08348 Cabrils, Spain
[5]USDA-ARS, Plant Genetic Resources Unit, Cornell University, Geneva, NY 14456, USA
[6]Department of Horticulture, Clemson University, Clemson, SC 29634, USA;
[7]Department of Horticulture, University of Arkansas, Fayetteville, AR 72701, USA
[8]Department of Horticultural Sciences, Texas A&M, College Station, TX 77843, USA
[9]Department of Plant Science, University of California - Davis, Davis, CA 95616, USA
[10]Department of Horticulture & Landscape Architecture, Washington State University, Pullman, WA 99164, USA
[11]Istituto di Genomica Applicata, Parco Scientifico e Tecnologico, 33100 Udine, Italy
[12]Department of Botany and Plant Pathology, OSU, Corvallis, OR 97331
[13]Department of Horticulture, Michigan State University, East Lansing, MI 48824, USA

ABSTRACT

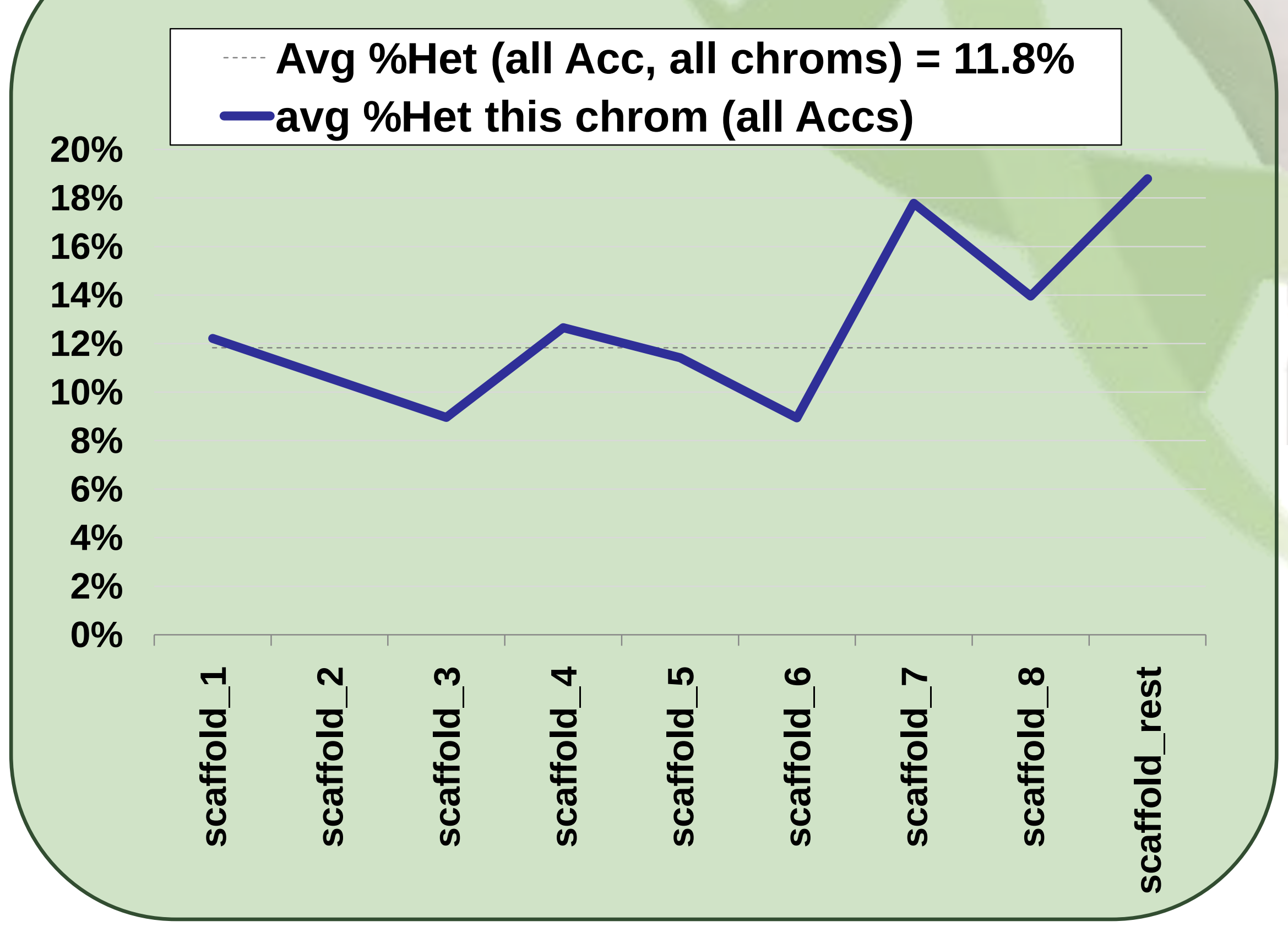
In the USDA-funded multi-institutional and trans-disciplinary project, “RosBREED”, crop-specific SNP genome scan platforms are being developed for peach, apple, strawberry, and cherry at a resolution of at least one polymorphic SNP marker every 5 cM in any random cross, for use in Pedigree-Based Analysis. RosBREED focuses on fruit quality traits. The majority of genes targeted for SNP development will be those shared among crops, and specific cases of synteny will be exploited to leverage ancestral relatedness within the Rosaceae family. In peach, we are re-sequencing genomic DNA from about 60 peach accessions coordinated between the U.S. and the Italian Consortium for Peach Genome Sequencing (DRUPOMICS project). These accessions were chosen for worldwide use in peach breeding and for diversity of genetic background, geographical origin, and phenotype. The peach genomic libraries are being sequenced with the Illumina Genome Analyzer IIx (GAIIx) at 3X coverage using paired-end reads (2x80bp) and indexed adaptors. The resulting peach SNP platform will be a valuable resource available to the worldwide peach research community.

Genome Coverage



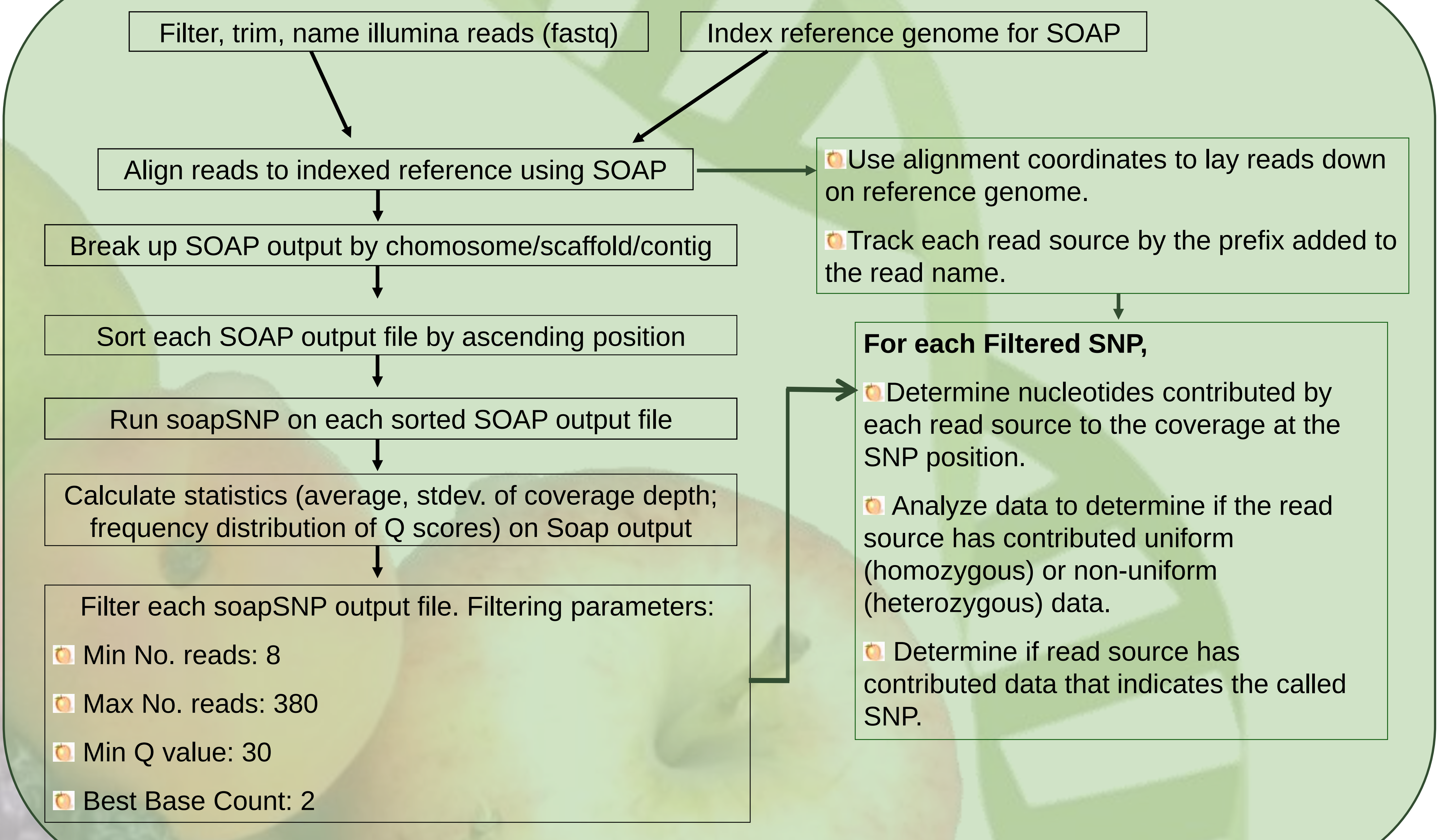
Genome coverage ranged from 0.20 in ‘Elberta’ to 5.2 in ‘Georgia Bell’. The number of SNPs per 10 Kb ranged from 0.15 in ‘Yellow St. John’ to 1.43 in ‘Georgia Bell’, as a function of genome coverage

Heterozygosity



Average percent heterozygosity was 11.8%
- Lowest in scaffolds 3 and 6
- Highest in scaffolds 7 and 8

SNP Calling Workflow (SOAP-based)



Filter, trim, name illumina reads (fastq)

Index reference genome for SOAP

Align reads to indexed reference using SOAP

Break up SOAP output by chomosome/scaffold/contig

Sort each SOAP output file by ascending position

Run soapSNP on each sorted SOAP output file

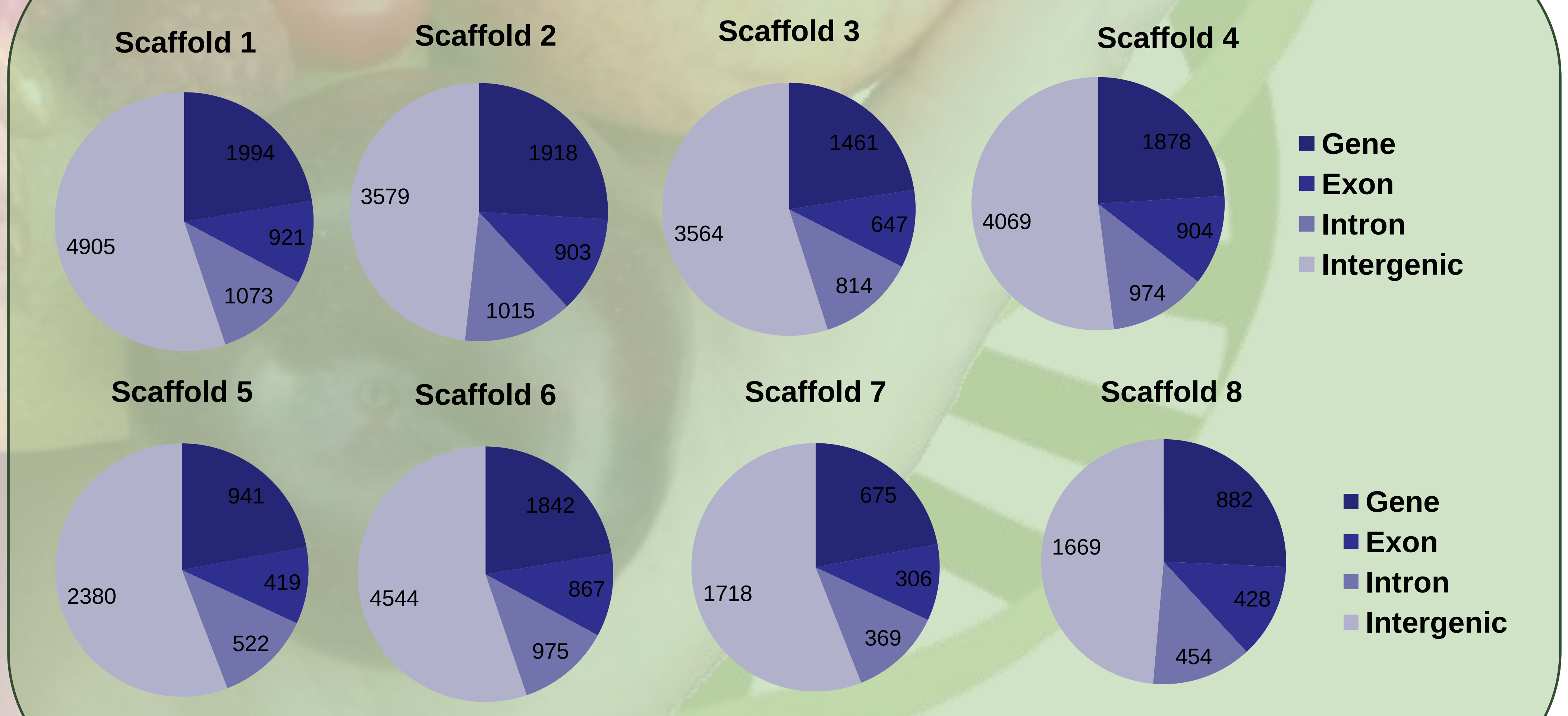
Calculate statistics (average, stdev. of coverage depth; frequency distribution of Q scores) on Soap output

Filter each soapSNP output file. Filtering parameters:
- Min No. reads: 8
- Max No. reads: 380
- Min Q value: 30
- Best Base Count: 2

- Use alignment coordinates to lay reads down on reference genome.
- Track each read source by the prefix added to the read name.

For each Filtered SNP,
- Determine nucleotides contributed by each read source to the coverage at the SNP position.
- Analyze data to determine if the read source has contributed uniform (homozygous) or non-uniform (heterozygous) data.
- Determine if read source has contributed data that indicates the called SNP.

No. and Location of SNPs in each Scaffold



CONCLUSIONS

SNP identification using SOAP in peach identified 260,181 SNPs in 23 genotypes using these filtering parameters. Up to 47,939 of these SNPs were heterozygous. The number of SNPs was mostly proportional to coverage depth. However, it this approach generated more than enough SNPs for genome scan. Ninety-six SNPs were chosen to validate this approach using the GoldenGate genotyping assay. A SNP consortium will use the resulting information to develop a list of SNPs to order fro the Illumina Infinium platform. The resulting peach SNP platform will be a valuable resource available to the worldwide peach research community.

Funding for *RosBREED: Enabling marker-assisted breeding in Rosaceae* is provided by the Specialty Crops Research Initiative Competitive Grant 2009-51181-05808 of the USDA's National Institute of Food and Agriculture.