Medicago truncatula Resequencing of 384 Lines

Joann Mudge

National Center for Genome Resources

NCGR Sequencing Center

8 solexa sequencing machines

illumina® CSPro™ CERTIFIED SERVICE PROVIDER

Over 350 runs

1st Certified Service Provider in North America

On-site full-time Illumina Field Service Engineer

NCGR

Genome Center IT Infrastructure

- Data Storage
  - Sun X4540 - Dual Processor, 2 X 4 Core, 32GB RAM, 48TB file server - hoover
  - Sun X4500 - Dual Processor, 2 X 2 Core, 32GB RAM, 24TB file server - silo
- Image processing/Basecalling
  - Sun X4150 – 2 x 4-core processors, 8GB RAM, 584GB disk– madrid, shiprock
- Alignments, Variant detection and/or expression analysis - Alpheus
  - 50 Sun x6220 Blades (200 Cores, 800G RAM, 14.6 TB disk) – shredder
  - 30 x Dell Blades (60 hyperthreaded cores 120GB RAM, 2.1TB disk)
  - 2 x Sun X4450 – 4 x 4-core processors, 64GB RAM , 584GB disk – aztec, tesuque
  - 1 x Sun X4600 – 2 x 2-core processors, 64GB RAM , 584GB disk – lasvegas
- Sybase DB storage - Alpheus
  - 2 x Sun V490 - 4 x 2-core SPARC processors, 16GB RAM, 292GB disk
  - 6140 SAN with 40 TB disk, 4 Gb/s Switch fabric
- Wx2 DB storage – Alpheus 9TB DB Storage
  - 4 x Sun X4240 – 4 x 4 core processors, 64GB RAM, 2336GB disk
  - 1 x Sun X4140 – 4 x 4 core processors, 64GB RAM, 1168GB disk

NCGR

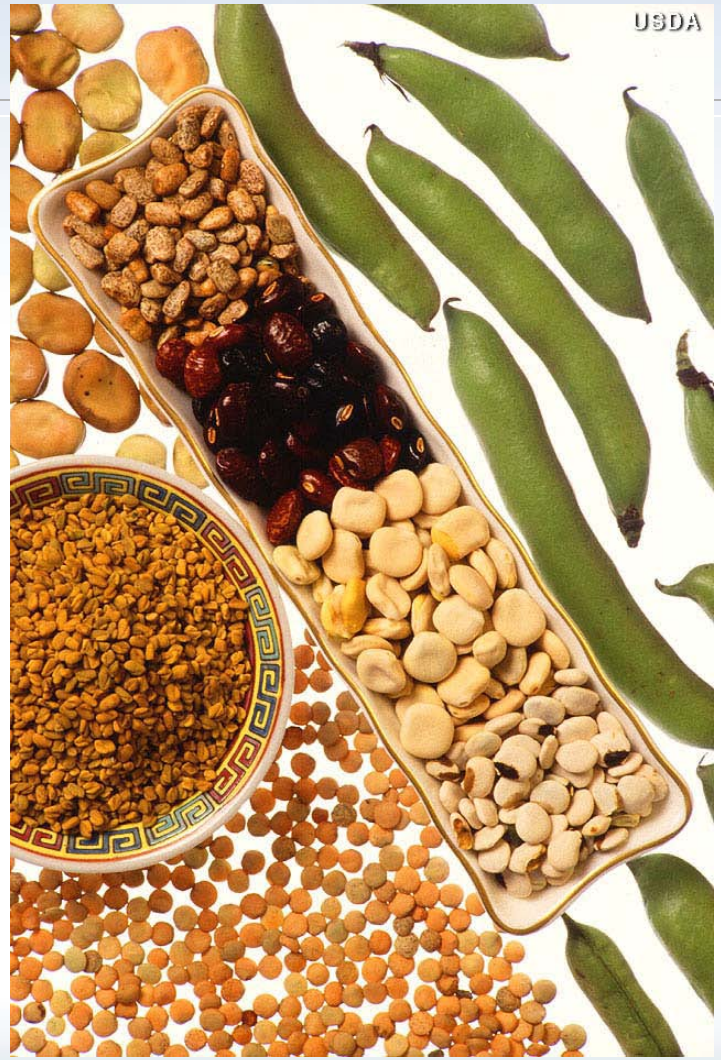
Genome Center IT Infrastructure

	Cores	RAM	Disk
Data Storage	12	64 GB	72 TB
Image processing/Basecalling	8	16 GB	1 TB
Alignments, Variant detection, Gene expression - Alpheus	296	312 GB	3.5 TB
Sybase DB storage - Alpheus	16	32 GB	41 TB
Wx2 DB storage – Alpheus 9TB DB Storage	80	320 GB	10.5 TB

NCGR

Medicago truncatula

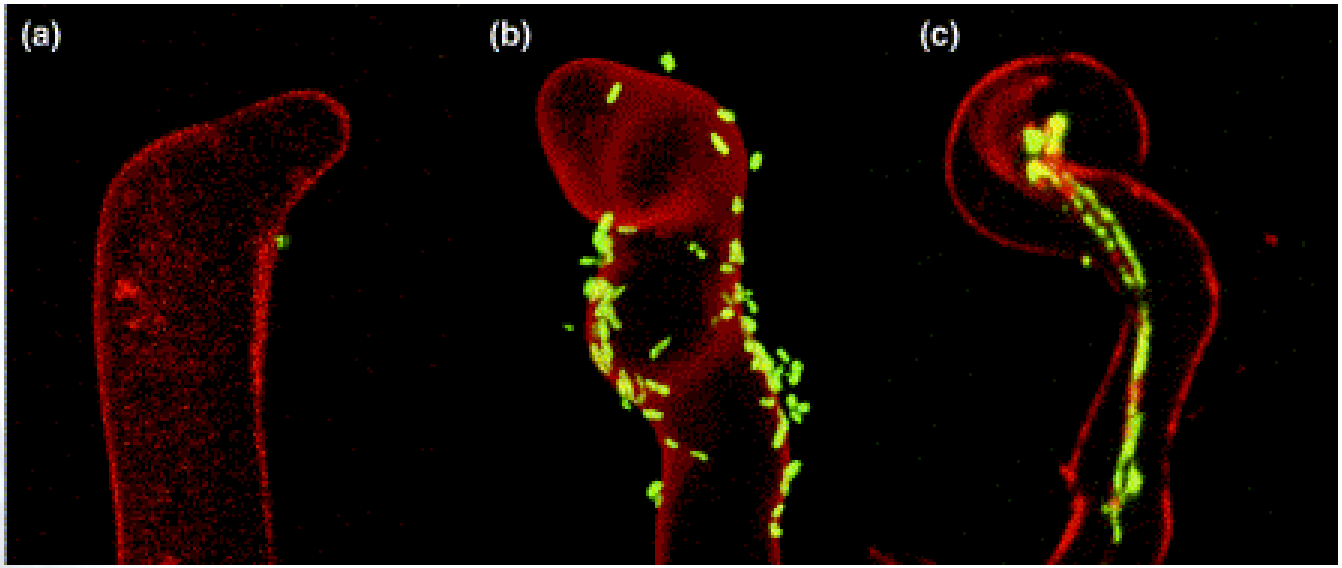
Model legume



NCGR

Medicago truncatula

Model for Symbiosis and Nitrogen Fixation



Limpens & Bisseling (2003) Curr. Opin. Plant Biol. 6: 343

NCGR

Medicago truncatula

- BAC by BAC reference assembly of A17
  - Euchromatic chromosome arms
  - Approximately 300 Mb of 500-550 Mb genome
  - International Consortium

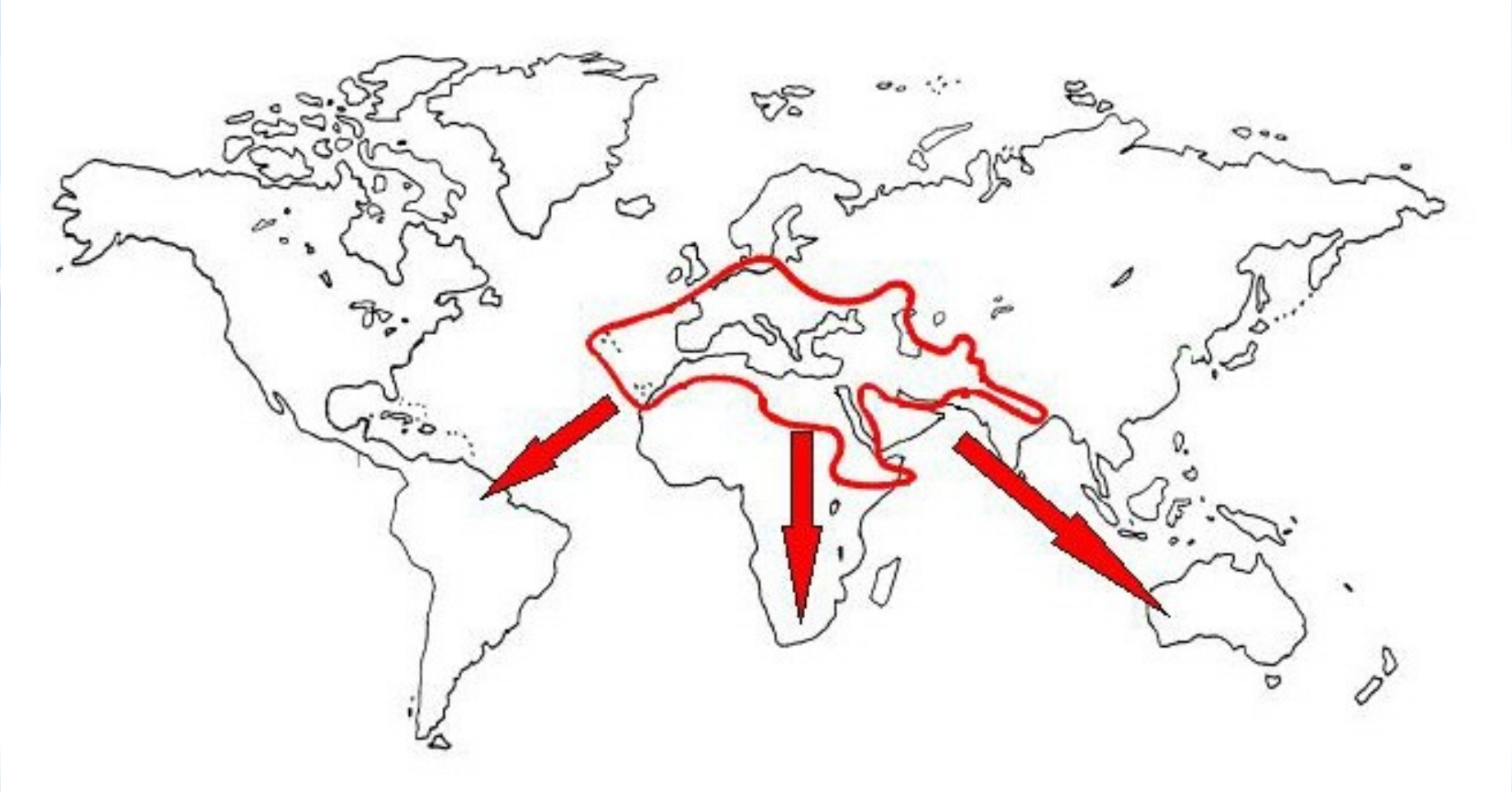
Medicago Hapmap Project

- Sequence 30 lines deeply
  - ~30X coverage
- Sequence 350-400 additional lines
  - ~5X coverage
- Variant detection
  - SNP
  - Indel
  - Structural variant
- SNP imputation
- Haplotype definition
- Population genetic studies
- Comparisons to symbiotic phenotypes

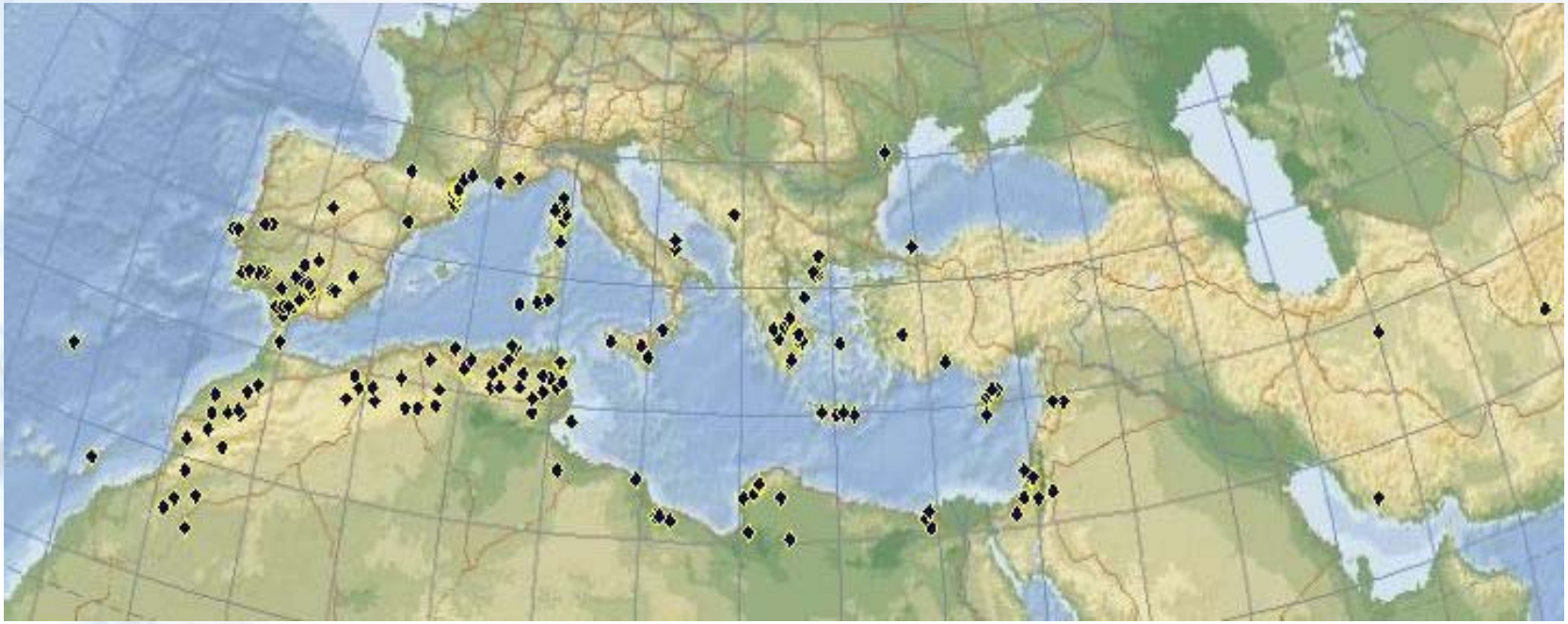
Currently we have sequence data for 18 lines

NCGR

Origin and distribution



Core Collections

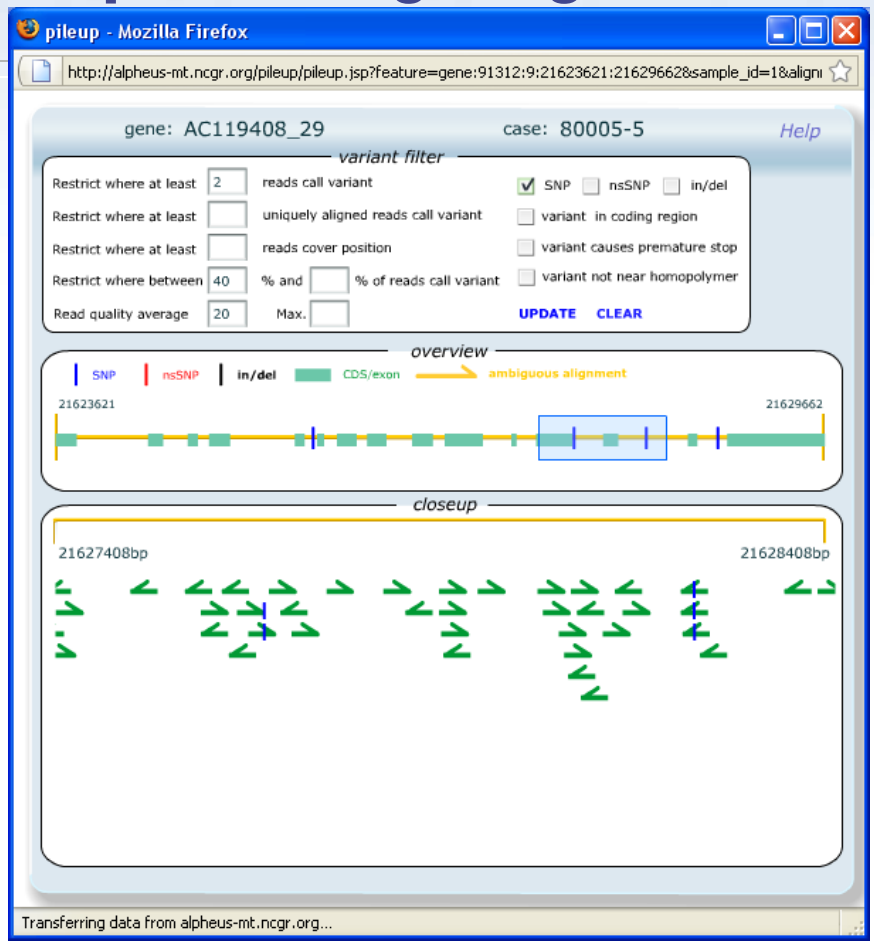


http://www1.montpellier.inra.fr/BRC-MTR/mauguio/mau

Strategy

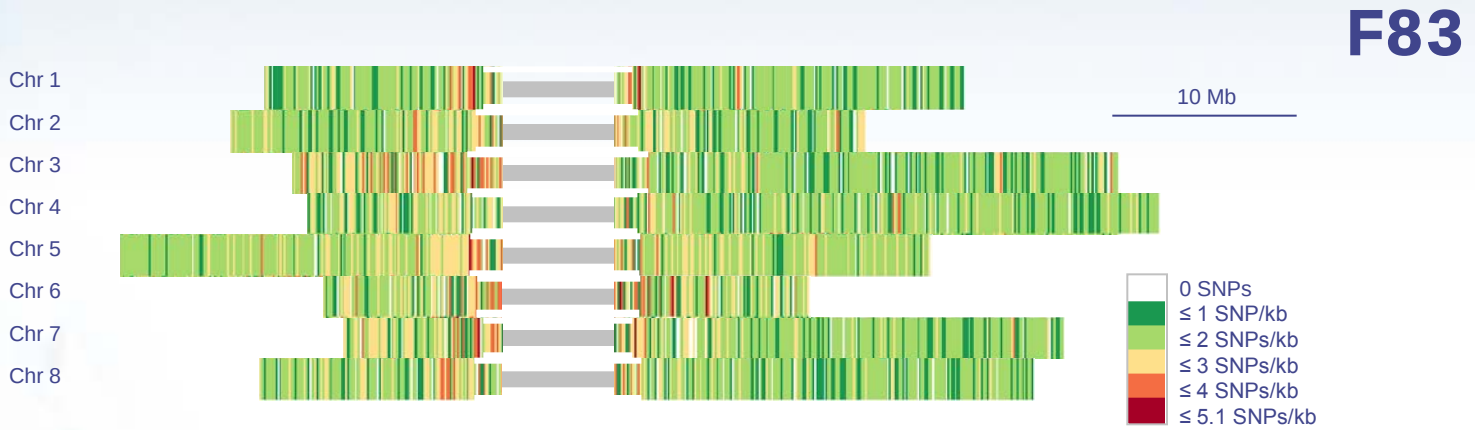
1. Solexa Sequencing
2. Alignment (GSNAP)
3. Variant Detection (Alpheus)

alpheus.ncgr.org



NCGR

Skim sequencing (1.5X coverage)



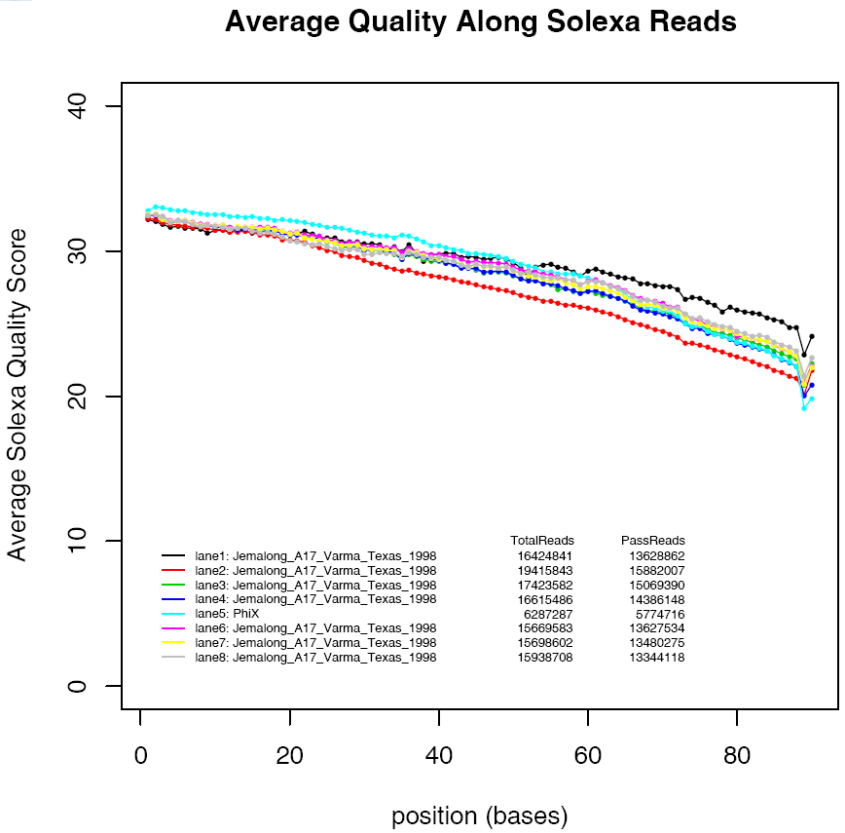
1.5x coverage

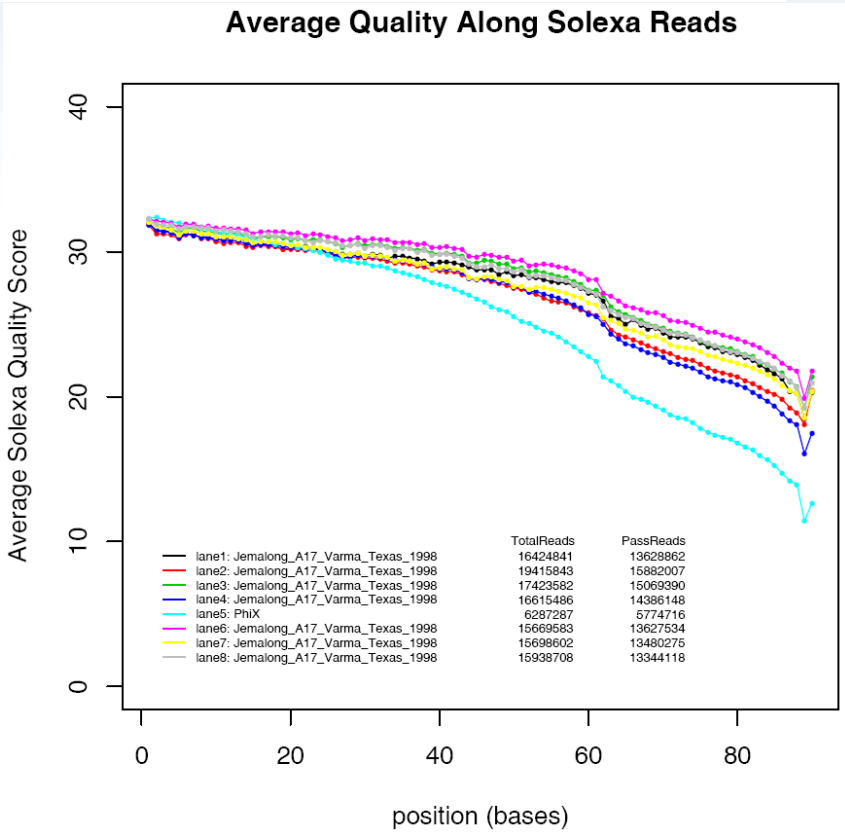
1/3 of genome has enough coverage for snp detection

160k SNPs

5 SNPs per kb

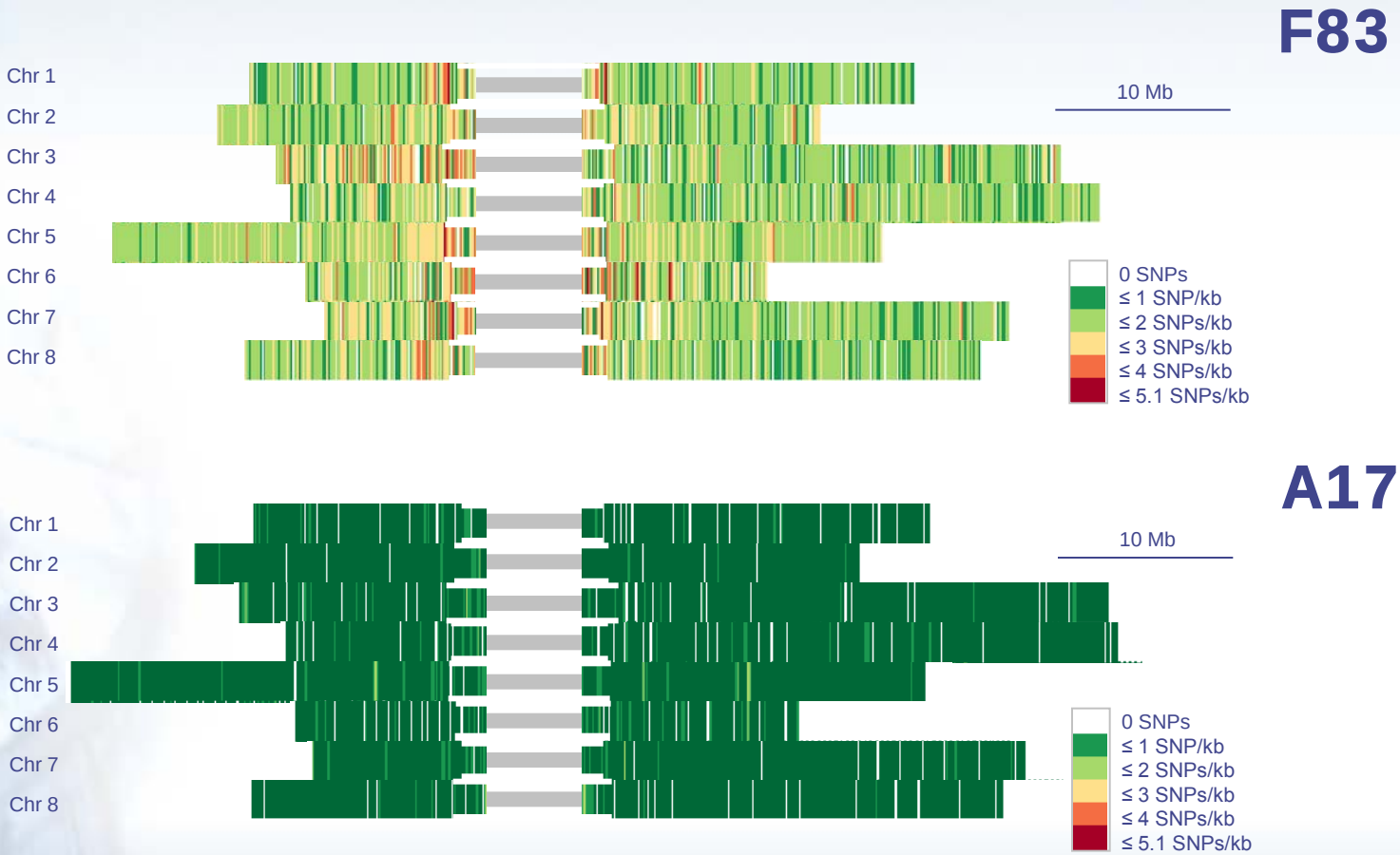
Single run: 18 Gb and 35X coverage



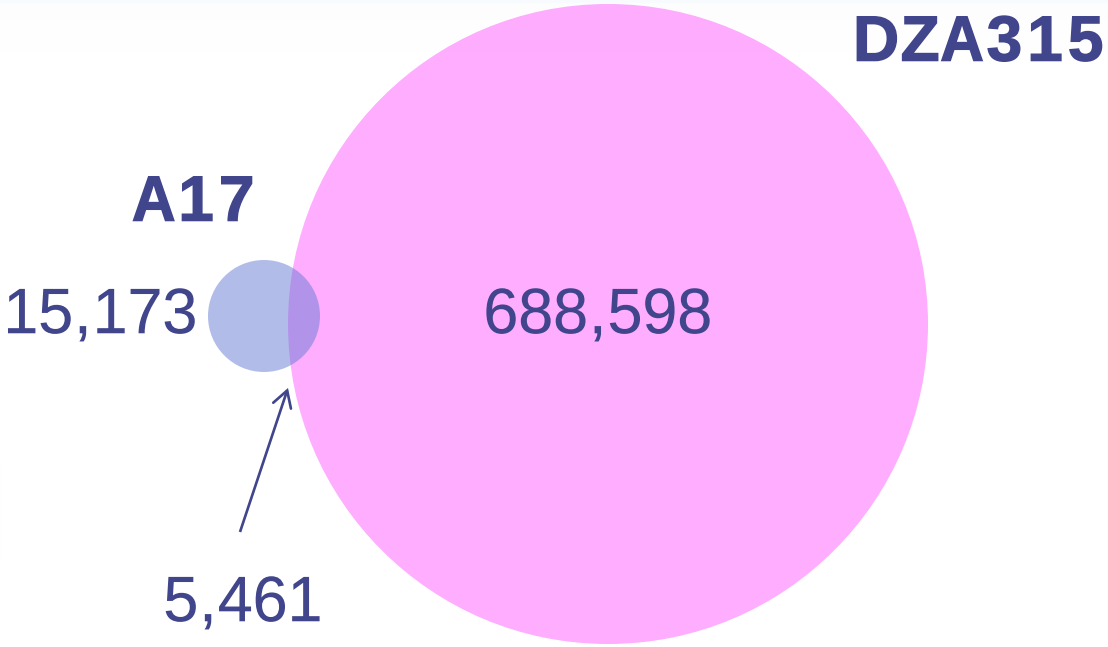


NCGR

Resequencing of A17



Venn Diagram of SNPs



Acknowledgements

NCGR

Greg May
Andrew Farmer
Ryan Kim
Jimmy Woodward
Leonda Clendenen
Jenny van Velkinburgh
Ray Langley
Darrell Diniwiddie

UMN

Nevin Young
Michael Sadowsky
Peter Tiffin

Boyce Thompson Institute

Maria Harrison

Hamline University

Betsy Martinez-Vaz

Plant Science
-Program Status-

- ***Medicago* HapMap*- Mudge & May**
  - National Science Foundation
  - University of Minnesota, Boyce Thompson Inst.
- ***Cacao* (Chocolate) Genome***
  - USDA-ARS & Mars, Inc.
  - Washing State University, JGI, USDA-ARS, IBM, Mars, Inc., PIPRA, CUGI
- **Sorghum Transcriptome - Kim**
  - USDA-ARS
- **Cotton Genome**
  - Texas Tech University
- **Legume Disease Resistance**
  - National Science Foundation
  - University of California – Davis
- **Chickpea & Pigeon Pea Diversity**
  - CIMMYT - Generation Challenge Program
  - ICRISAT
- **Legume Information System‡**
  - **USDA-ARS**
  - **Iowa State University**
- **LegumeDB – Retzel**
  - **USDA-ARS**
- **Semantic Web* – Retzel**
  - **National Science Foundation**
  - **University of Arizona**
- **Soybean Rust**
  - **UDSA-ARS**
  - **Iowa State University**
- **Genetic Diversity in Maize‡**
  - **Syngenta**
- **Homologous Recombination Genes* - Retzel**
  - **BRDC**
  - **University of Minnesota**
- **Johne's Disease* - Retzel**
  - **BRDC**
  - **University of Minnesota**

PI & CoPI Institutions and Sponsors

(*New or ‡ renewed award for 2009)