

# SNP Diversity in Cultivated Tomato

Joanne A. Labate

Larry D. Robertson

Angela M. Baldo

USDA-ARS Plant Genetic Resources Unit  
Geneva, NY



## Recent diversity studies of cultivated tomato

fresh market, processing, heirloom lines (Van Deynze et al. 2007)

European commercial greenhouse cultivars (van Berloo et al. 2008)

Italian landraces, vintage, modern cultivars (Mazzucato et al. 2008)

Greek landraces, cultivars (Terzopoulos and Bebeli 2008)

cultivars, hybrids, elite lines (Chen et al. 2009)

landraces from centers of diversity, cultivars (Labate et al. 2009)

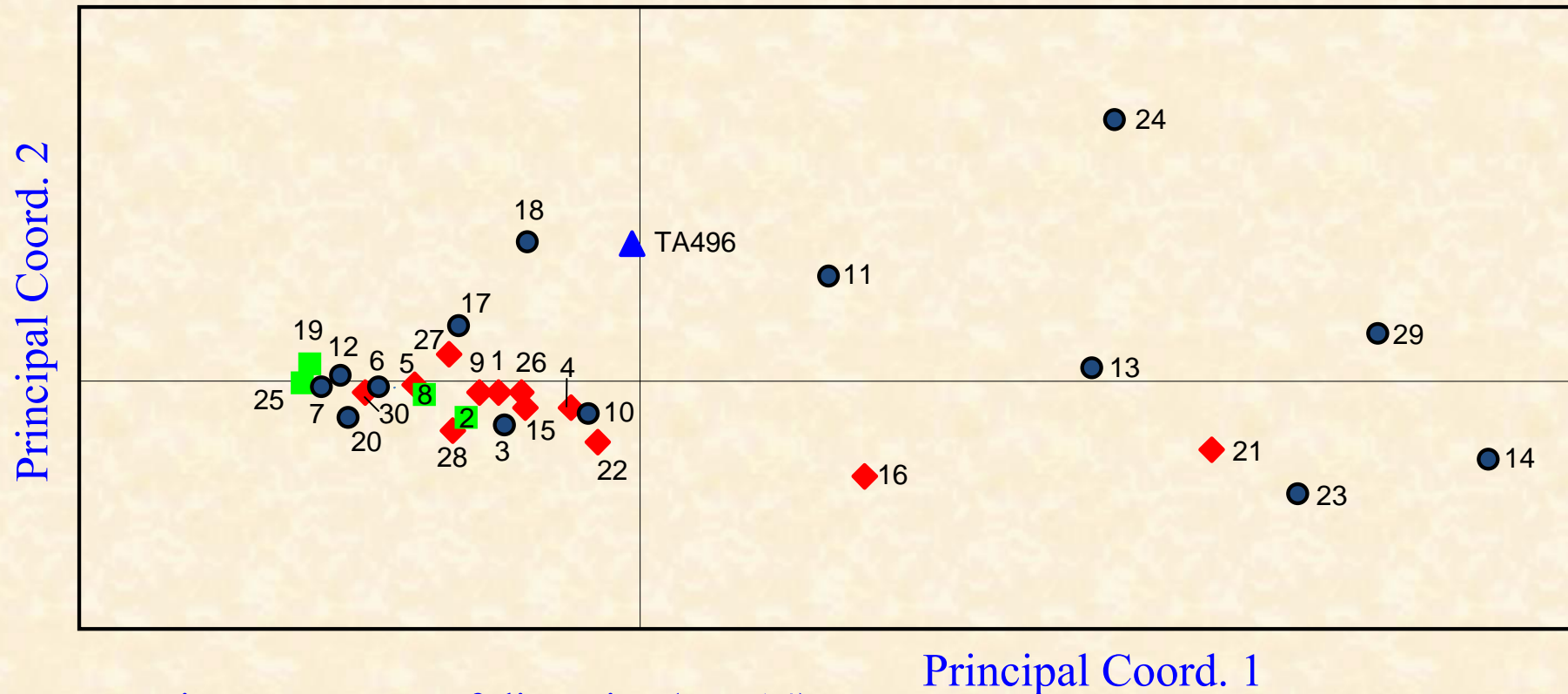
- discover polymorphic markers that may be transferable to elite germplasm
- utilize unimproved, cultivated germplasm as sources of new alleles

## PGRU Diversity Panels (30-50 NPGS accessions)

- DNA sequences 50 loci (gene fragments)
- SNP genotypes (D. Francis lab)
- Agronomic traits and fruit quality (2008, 2009)



Genetic diversity among 31 tomato lines based on 50 sequenced markers  
via covariance matrix with data standardization



- primary centers of diversity (n = 14)
- secondary centers of diversity (n = 4)
- ◆ countries contiguous with primary centers (n = 12)
- ▲ line TA496

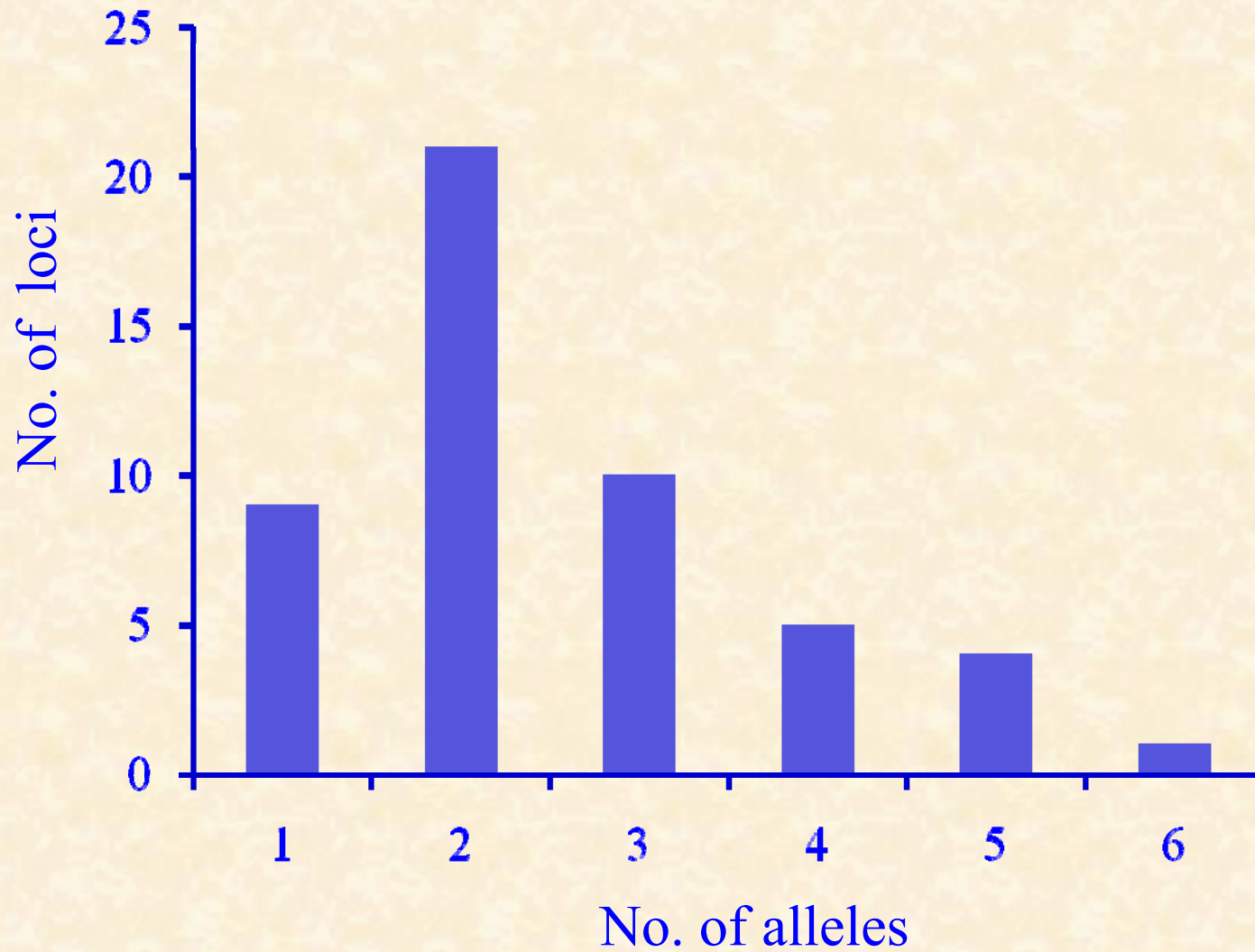
Mean diversity across 50 loci for three *S. lycopersicum* samples

Population	no. <sup>a</sup>	no. of alleles	<i>F</i>	<i>H<sub>e</sub></i> <sup>b</sup>
Primary center	14	2.3	0.793	0.254
Contiguous w/ primary center	12	1.7	0.947	0.152
Secondary centers	6	1.2	1.000	0.092

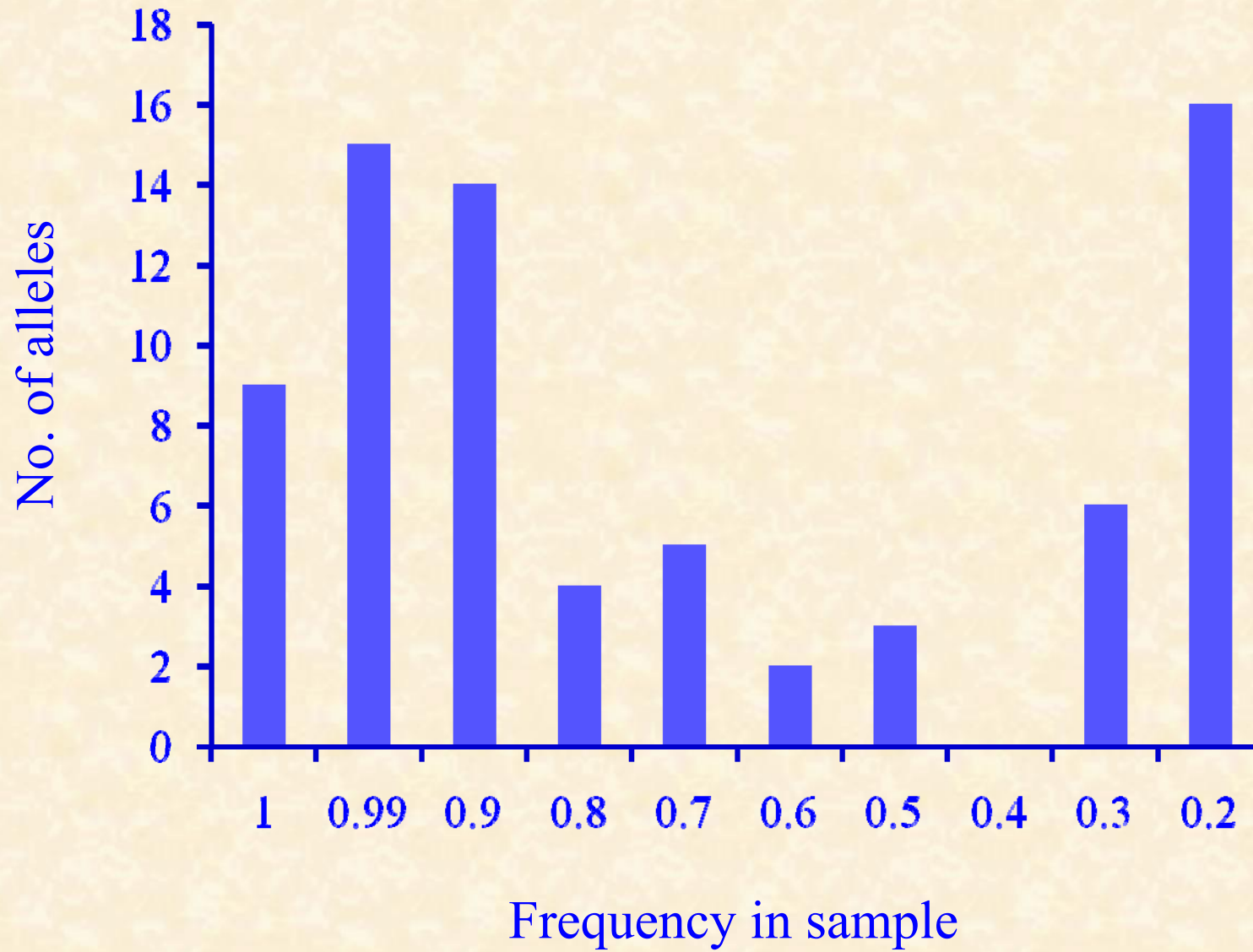
<sup>a</sup>Number of plants sampled

<sup>b</sup>Expected heterozygosity =  $2n(1-\sum x^2)/(2n-1)$  where  $n$  is the number of sampled chromosomes,  $x$  is allele frequency; each haplotype was treated as an allele

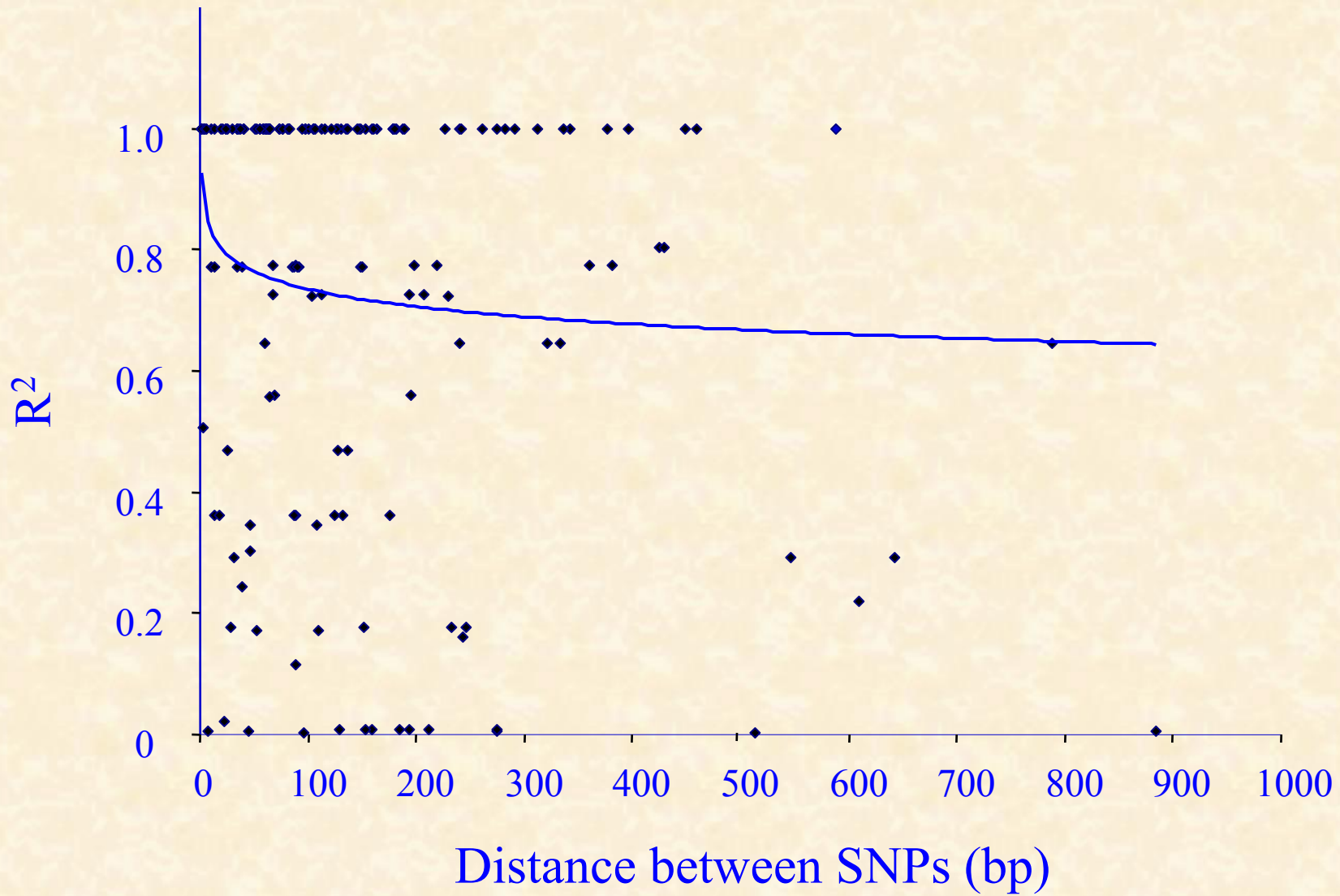
Number of alleles per locus in *S. lycopersicum*, 50 loci



Allele frequency in *S. lycopersicum*, 50 loci



# Within-locus linkage disequilibrium in *S. lycopersicum*, 21 loci





## Between-locus linkage disequilibrium in *S. lycopersicum*

41 loci, 797 pairwise comparisons:

- 152 pairs (0.191) in LD ( $P \leq 0.05$ )
- for 75/152 pairs with known map locations, only 8 pairs (0.107) were located on same chromosome

## Diversity estimates for various crops and wild relatives

---

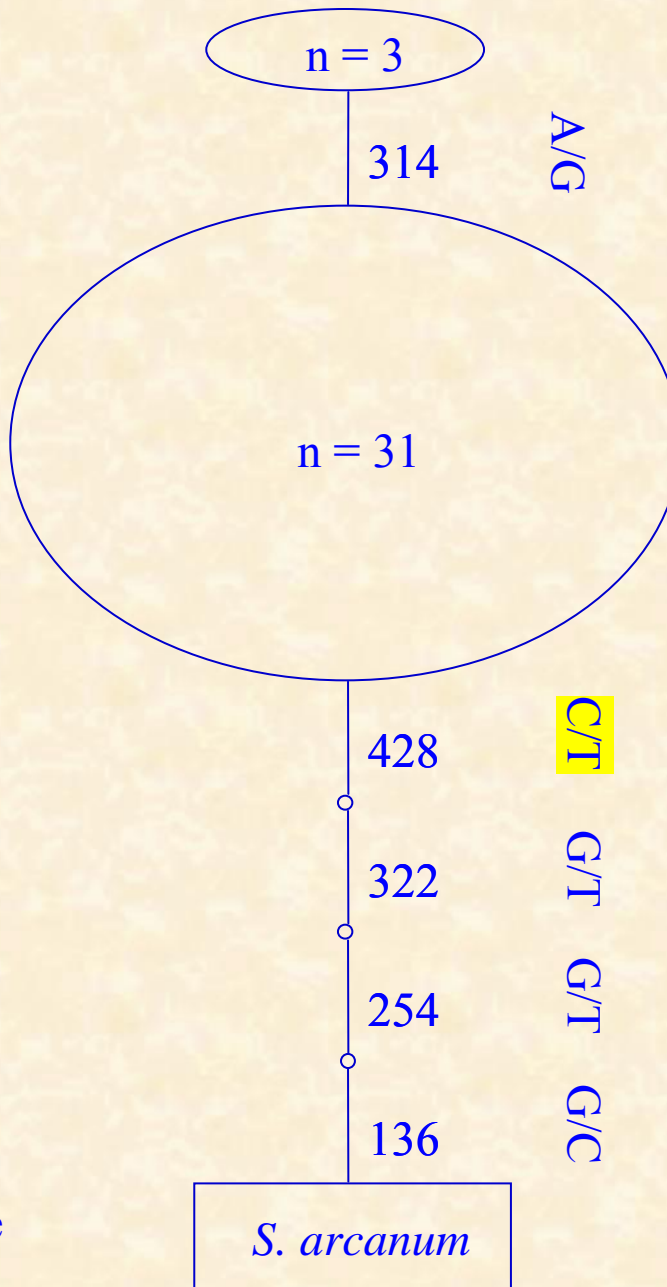
Species	Mating system <sup>a</sup>	SNPs/kb <sup>b</sup>	Reference
Soybean	S	1.1	Zhu <i>et al.</i> 2003
Tomato	S	1.3	Labate <i>et al.</i> 2009
Sorghum	S	2.2	Hamblin <i>et al.</i> 2006
Asian rice	S	2.3	Caicedo <i>et al.</i> 2007
Wild barley	S	7.5	Morrell <i>et al.</i> 2005
<i>S. chilense</i>	SI	11.0	Arunyawat <i>et al.</i> 2007
<i>S. peruvianum</i>	SI	12.9	Arunyawat <i>et al.</i> 2007

---

<sup>a</sup>S, self-fertilizing; SI, self-incompatible

<sup>b</sup>Expected heterozygosity for a pair of SNPs sampled at random

Haplotype network  
C2\_At1g32130  
(430 bp)

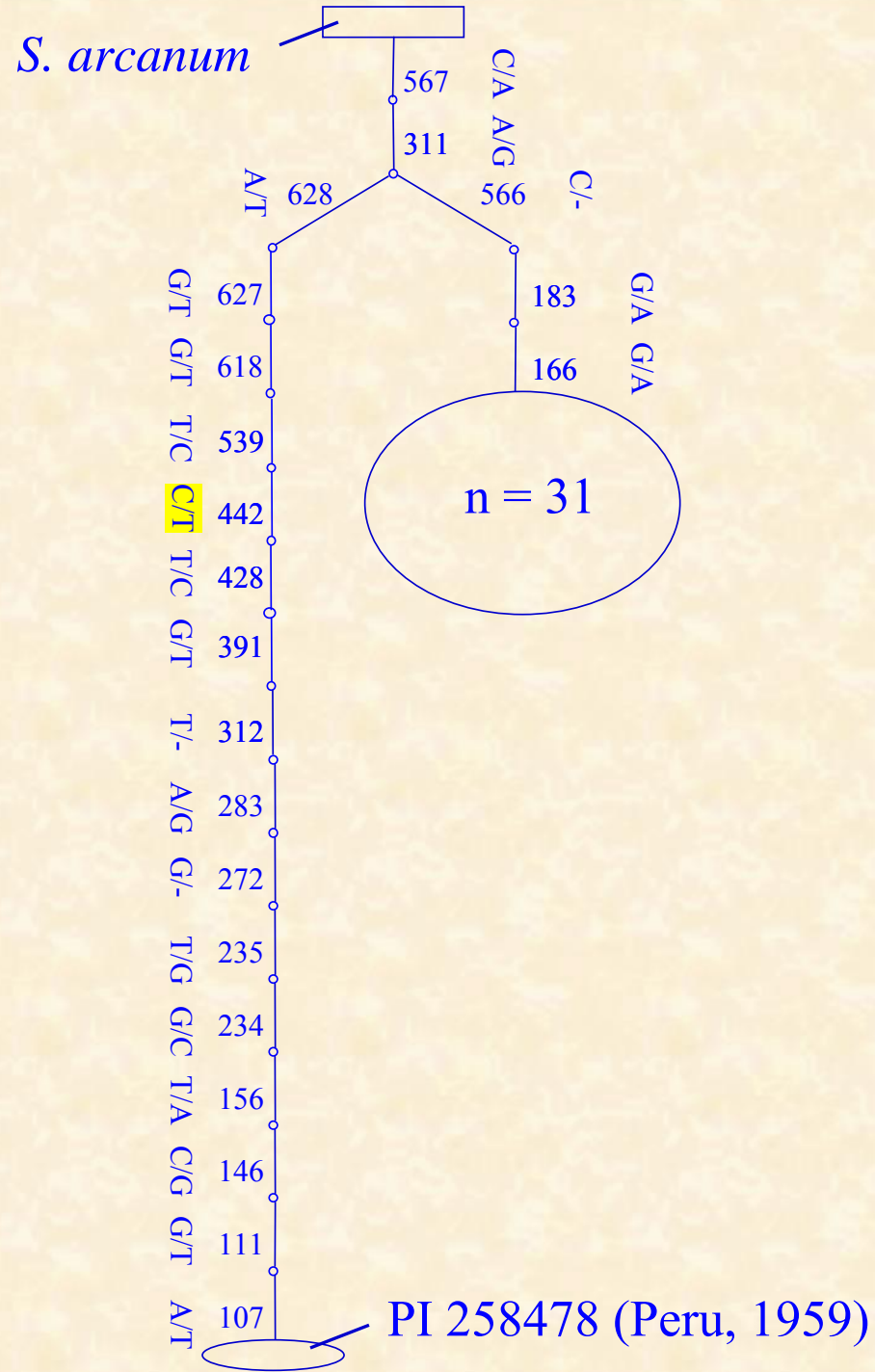


- noncoding (no shading)
- synonymous (shaded yellow)
- non-syn, conservative (shaded cyan)
- non-syn, nonconservative (shaded pink)

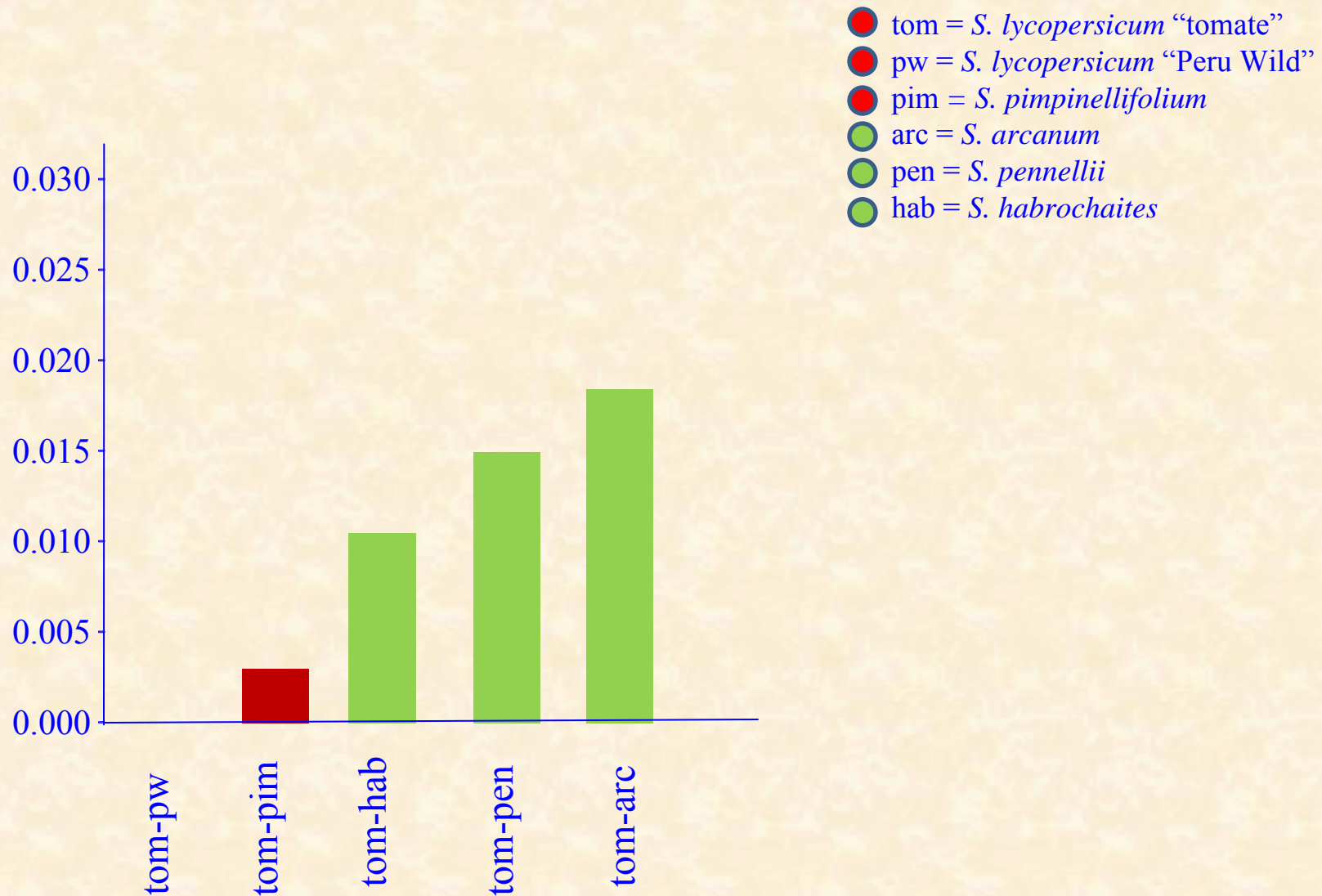
Haplotype network  
2819\_5  
(656 bp)

“hypothetical introgression”

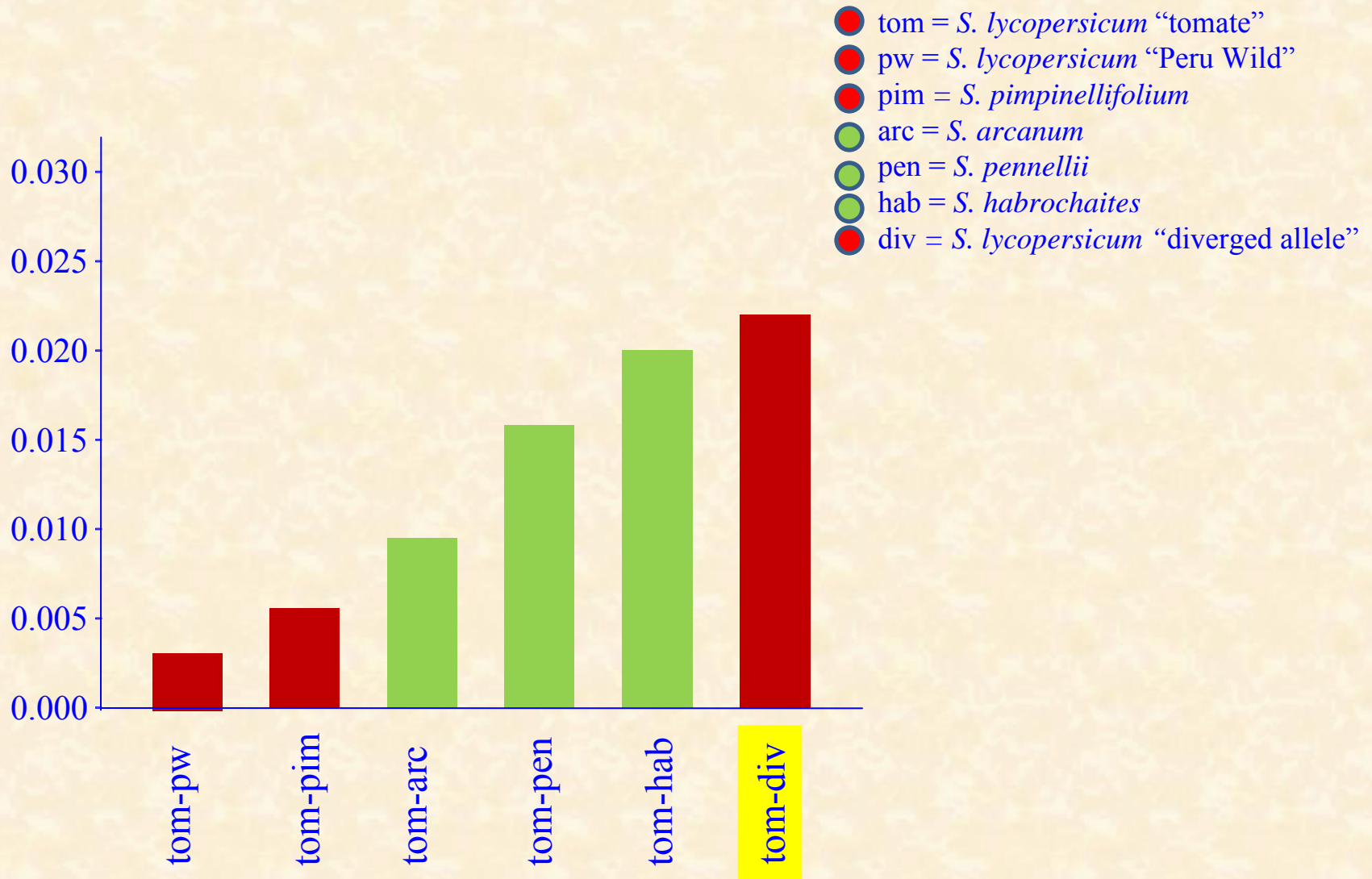
- noncoding (no shading)
- synonymous
- non-syn, conservative
- non-syn, nonconservative



## Mean pairwise divergence between sequences at two control loci



# Mean pairwise divergence between sequences at five loci with hypothetical introgressions



## Summary and conclusions

- diversity within *S. lycopersicum* similar to other selfing crops (1 – 2 SNPs/kb)
- most loci (80%) carry at least two alleles
- variant alleles are rare in sample ( $\leq 10\%$ )
- individuals are highly homozygous ( $F = 0.8 - 1.0$ )
- low to moderate population diversity ( $H_e = 0.09 - 0.25$ )
- high within-locus LD
- moderate to low between-locus LD
- highly diverged haplotypes are observed  $\sim 10\%$  (cryptic introgressions?) Islam et al. 2004

Lab

S. Sheffer  
A. Michaels  
K. Timmer

Bioinformatics

W. Lamboy  
M. D'Amico

Seed Crops

P. Kisly  
D. Sharman  
J. Oughterson  
B. Garman  
T. Balch

Collaborators

S. Tanksley, F. Wu, L. Mueller, R. Ahrens, Cornell Univ  
J. Giovannoni, USDA-ARS  
D. Spooner, USDA-ARS  
D. Francis, The Ohio State University

*Thank you!*



Funded by CRIS project 1910-21000-019-00D