Welcome to the Development of SNP-Based Tetraploid Maps for Potato Webinar

Presentation Available at: http://www.extension.org/pages/63187

Today's Presenters: Dr. Dave Douches & Joe Coombs

Brought to you by: The Plant Breeding and Genomics Community of Practice

Host: Heather Merk

Sign up for PBG News: http://pbgworks.org



Please fill out the survey evaluation: (You will be contacted via email)

Watch past webinars and sign up for future webinars! http://www.extension.org/pages/60426

Development of SNP-Based Tetraploid Maps for Potato



David Douches¹, Kim Felcher¹, Joseph Coombs¹, Dan Zarka¹, Allen Van Deynze², John Hamilton¹, Candice Hansey¹, C. Robin Buell¹ ¹Michigan State University, East Lansing, MI 48824

²University of California-Davis, CA 95616

Objectives

- Visualizing SNPs in Genome Studio including Five Cluster Calling
 Filtering SNPs for tetraploid mapping populations
- Generating a map using TetraploidMap software
- Initiating QTL analysis
 - Single Marker ANOVAs
 - TetraploidMap



Genetic maps in potato

Several linkage maps have been constructed for potato

- (Bonierbale et al. 1988; Gebhardt et al. 1991; Jacobs et al. 1995; van Os et al. 2006, Felcher et al. 2012 (accepted))
- Diploid potato populations are often used for linkage mapping
- Map size ranges from 606 cM to 1120 cM
- Genetic markers per map range from 85 to 10,000 markers
- Markers used include isozymes, RFLPs, SSRs, AFLPs and more recently SNPs



SolCAP Genome-wide set of SNP markers: potato activities

Assess concordance between map location of SNPs and the potato genome sequence location (based on pseudomolecules).

- Felcher et al. 2012 (PLoS ONE, accepted)

- Use the Infinium 8303 potato array to genotype tetraploid mapping populations and generate SNP-based genetic maps
- QTL Analysis of tetraploid populations
- Association analysis of potato diversity panel



Integration of Two Diploid Potato Linkage Maps with the Potato Genome Sequence

DRH (92 progeny)

 DM x RH (from Virginia Tech) was selected for mapping because the RH parent has been used extensively in potato mapping studies and genome sequencing.

D84 (92 progeny)

 DM x 84SD22 (from MSU) was selected for mapping because 84SD22 was shown to have a higher percentage of polymorphic SNPs.





D84 Chromosomes 1-6



D84 Chromosomes 7-12

7	8	9	10	11	12
0.0 1.2 1.2 1.2 1.2 1.2 1.2 1.2 1.2	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	$\begin{array}{c} 0.0 \\ 2.5 \\ 3.8 \\ c1 \\ c12239 \\ c2 \\ 41766 \\ c2 \\ 28761 \\ c2 \\ 4154 \\ c1 \\ 9058 \\ 318 \\ 318 \\ c1 \\ c1 \\ 9058 \\ 318 \\ c1 \\ c1 \\ 9058 \\ 318 \\ c1 \\ c1 \\ 9058 \\ c2 \\ 44154 \\ c1 \\ 9058 \\ c2 \\ 3186 \\ c1 \\ 1590 \\ c2 \\ 24409 \\ c1 \\ 15590 \\ c2 \\ 27829 \\ 455 \\ c1 \\ 16590 \\ c2 \\ 27829 \\ 455 \\ c1 \\ 16590 \\ c2 \\ 27829 \\ c1 \\ 1590 \\ c1 \\ 12584 \\ 503 \\ c2 \\ 2831 \\ 51.0 \\ c1 \\ 13006 \\ c2 \\ 42429 \\ 53.5 \\ c2 \\ 1335 \\ c2 \\ c2 \\ c2 \\ c3 \\ c2 \\ c2 \\ c2 \\ c2$	$\begin{array}{c} 0.0 \\ 1.3 \\ 2.6 \\ 2.6 \\ 2.6 \\ 2.6 \\ 2.7 \\ 2.6 \\ 2.6 \\ 2.7 \\ 2.6 \\ 2.6 \\ 2.7 \\ 2.6 \\ 2.6 \\ 2.7 \\ 2.7 \\ 2.6 \\ 2.7 \\ 2.7 \\ 2.6 \\ 2.6 \\ 2.7 \\ 2.7 \\ 2.6 \\ 2.6 \\ 2.7 \\$	$\begin{array}{c} 0.0\\ 3.9\\ 1c1 \\ 7744 \\ c2 \\ 24595 \\ c1 \\ 7744 \\ c2 \\ 15641 \\ c2 \\ 1624595 \\ c1 \\ 1724 \\ c2 \\ 1725 \\ c2 \\ 1105 \\ c2 \\ c2 \\ 1105 \\ c2 \\ c2 \\ 1105 \\ c2 \\ c$

<u>MICHIGAN STATE</u> UNIVERSITY

Comparison of SNPs in the 2x populations DRH and D84

			Number		Length					
	Includes	Co-segre	egating SNPs	Марр	ed Segre	gating SNPs	(cl	M)	(№	1b)
Chromosome	DRH	D84	Common	DRH	D84	Common	DRH	D84	DRH	D84
1	268	279	114	121	76	14	125	98	81	81
2	208	270	103	97	55	17	79	53	46	46
3	88	239	26	64	46	6	78	61	48	48
4	230	186	74	105	53	12	89	91	64	64
5	144	158	52	55	46	9	100	65	47	47
6	213	216	110	90	59	19	66	65	52	55
7	146	245	52	66	49	5	70	47	53	53
8	147	183	57	74	48	11	71	67	43	43
9	164	195	62	89	57	8	100	69	53	53
10	115	131	51	66	43	14	82	63	52	52
11	131	171	45	74	50	8	76	48	42	42
12	106	181	41	43	55	4	31	65	54	59
Total	1960	2454	787	944	637	127	965	792	634	642

Total mapped SNPs: 3627



Tetraploid Map Materials and Methods

- PRRG 184 Progeny
 - Premier Russet (PR) X Rio Grande Russet (RG)
 - USDA/ARS -- ID (Rich Novy)
- Phenotypic data
 - 21 Traits



- 3 locations (ID Novy, NC Yencho, MN Thill)
- -2 years (2010, 2011)



Traits being evaluated within SolCAP 10 plant plots, two replicates

- specific gravity
- chip color after cold storage
- sucrose/glucose
- skin texture
- tuber shape (l/w/h)
- eye depth
- skin color, flower color
- flesh color
- vine maturity (95, 120 dap)
- growth habit (prostrate, erect, etc.)
- total yield
- heat sprouts
- internal defects

"The key three"







SolCAP Germplasm & Populations

Panel (325 clones)

- Top 50 North American varieties
- Historical varieties
- Advanced breeding clones
 - from every US and Canadian program
- Non-American germplasm
- Genetic stocks
- 10 Wild species

Panel Analyses

- Association mapping
- Parental selection
- Resolve population structure

Russet Mapping population

- PR x RG (184 progeny)
- QTL analysis
- Diploid Mapping populations
 - DM x RH
 - DM x 84SD22



Other US Potato Mapping populations SNP genotyped

- Atlantic x Superior
 - (tuber calcium, reducing sugars, internal defects, specific gravity (starch))
- B1829-5 x Atlantic
 - (chip color, internal heat necrosis, specific gravity, maturity)
- Tundra x Kalkaska
 - (scab R, chip color, reducing sugars, specific gravity, asparagine, acrylamide)
- Jacqueline Lee x MSG227-2
 - (specific gravity, late blight resistance, vine maturity)
- Waneta x Pike
 - (specific gravity, chip color, disease resistance)
- W4 x 524-8 (diploid)
 - (specific gravity, chip color, disease resistance)



Potato RNA-Seq for SNP Development

Leaf

Flower

Callus

We sequenced the transcriptomes of three varieties using Illumina RNA-Seq.
Tuber

- Snowden (Chipping)
- Atlantic (Chipping)
- Premier Russet (Processing-FF)

2 lanes of 61 bp paired end reads for each variety.

Total Purity Filtered Sequence:

- Atlantic: 30,185,186 reads (1.84 Gb)
- Premier: 31,949,096 reads (1.94 Gb)
- Snowden: 33,288,120 reads (2.03 Gb)
- ~50-60x coverage of the transcriptome



Potato Illumina Infinium Platform Design Summary

69,011 SNPs passed all filtering steps

- Further filtering was performed to meet the design criteria for the Infinium platform using the following criteria:
 - Biallelic based on all available sequence
 - Within exons (mapped > 95% to DM1-3 draft genome sequence)
 - 50 bp from exon/intron junction
 - Max 1 SNP within 100 bp window of candidate SNP
 - Passed Illumina Scoring
 - 69,011 SNPs passed all filtering steps
 - All SNPs are on the potato genome browser

Infinium 8303 Potato Array Design

8303 SNPs in total

- 3018 SNPs from candidate genes
- 536 SNPs from previously mapped genetic markers
- 4749 dispersed SNPs selected to achieve maximum genome coverage



Infinium 8303 Potato Array (Single Base Extension)



- Unique oligo for each bead type
- Bead Pool is 250,000 per sample
- Random self-assembly of beads onto the chip
- Redundancy averages 15 to 30 beads of each type
- 8,303 SNPs on Illumina Infinium chip

CHICAN

24 samples per chip

Illumina iScan

The bead chips are loaded into the Illumina iScan.



- The iScan uses a laser to excite the red or green fluorophore of the single base extension product on the beads.
- The scanner records high-resolution images of the light emitted from the fluorophores.

 Data analyzed in Genome Studio software

Illumina Genome Studio Software

🏶 GenomeStudio - Genotyping - Diploid_mapping_population Window Analysis Tools Help 0 $\triangleleft \triangleright \mathbf{X}$ $4 \triangleright \mathbf{X}$ Full Data Table SNP Table Paired Sample Table SNP Graph Heat Map Ba 🚺 🔓 Project 🖑 🔍 🔎 🖾 🚺 彩 🖹 🕞 📭 🚑 蒜 部 🖉 💉 🦛 👖 🕥 📓 😥 💷 🔒 🚼 🚟 1 • ٠ solcap_snp_c1_10000 Sample 1 Sample 2 Sample DM 84SD22 1.40 GenTrain Index Address Chr Position Frac A Frac C Frac G Frac T R Theta R Score Th Name GType Score Theta GType Score GType 1.20 Score 0.275 0.108 solcap_snp... 4070. 0 0.6414 0.333 0.284 AA 0.4693 0.0491 1.2197 AB 0.4693 0.3948 1.5822 ΔΔ 0.4693 0.0 🗸 F solcap_snp. 0.7424 0.294 0.206 0.186 0.314 0.6494 0.9594 0.9417 0.6494 0.5091 1.2203 AB 0.6494 6777. 0 0 BB AB 0.5 - 2 œ 0.80 solcap snp. 0 0.6799 0.333 0.157 0.275 0.235 AA 0.600 0.0102 0.746 BB 0.6008 0.984r AB 0.6008 7368 0.5 Norm solcap_snp.. 5374. 0 0 0.8009 0.304 0.235 0.167 0.294 BB 0.7514 0.9729 1.1682 AB 0.7229 0.4121 1.5493 AB 0.7514 0.5 0.60 4769. 1.0154 0.9815 1.0283 solcap_snp... 0 Π 0.6883 0.196 0.167 0.275 0.363 AA 0.6169 0.0018 BB 0.6169 AB 0.6169 0.4 0.40 0.6935 0.284 0.294 0.6268 0.9933 1.0322 0.6268 0.0215 1.0586 0.6268 solcap snp. 1674 0 Π 0.196 0.225 BB AA AB 0.5 1.6023 1.4672 0.1994 solcap snp 4981 Π 0.4086 0.245 0.284 0.225 0.245 BB 0.1994 0.7267 AA 0.1994 0.0960 AF 0.4 0.20 0 0.304 0.245 0.255 1.0538 0.9179 solcap snp 6273 0 0.9176 0.196 BB 0.5295 0.9954 BB 0.5295 0.9853 BB 0.5295 0.9 solcap snp 4576 0 0.6853 0.304 0.235 0.176 0.284 AA 0.6113 0.0150 0.8187 BB 0.6113 0.9809 0.9824 AB 0.6113 0.5 22 10 5769 0 0.7677 0.265 0.235 0.235 0.265 BB 0.6945 0.9807 1.1252 AB 0.6945 0.5076 1.5619 AB 0.6945 0.5 solcap snp 0 -0.20 11 4970 0.7627 0.314 0.245 0.186 0.255 0.6857 0.9555 0.9181 0.6857 0.5100 1.2423 0.6857 0.4 0.40 0.60 solcap snp 0 BB AB AB 3772 0.9200 0.353 0.5327 1.0456 0.5327 1.0538 0.5327 Norm Theta solcap_snp 0 0.1370.216 0.294 AA 0.0044 AA 0.0020 AA 0.0 13 solcap_snp 7379 0 0.7716 0.235 0.186 0.363 0.216 AA 0.3559 0.1546 1.1397 ΔΔ 0.3559 0.1299 0.9964 0.3559 0.1 AA 14 solcap_snp 4366 0 0 0.6132 0.304 0.147 0.225 0.324 AA 0.4762 0.0341 0.9200 BB 0.4762 0.9538 0.9302 AB 0.4762 0.5 4 b X Samples Table 15 solcap_snp 3073 0 0.9024 0.265 0.284 0.196 0.255 AA 0.8929 0.0160 0.9812 AB 0.8929 0.4595 1.3713 AA 0.8929 0.0 0 16 6769 0 0.8535 0.490 0.147 0.167 0.196 BB 0.4485 0.9229 0.9701 NC 0.0057 1.0000 0.0147 BB 0.4485 0.9 solcap snp Π. 圖 🖗 🗎 🕞 🕒 💱 👬 🎊 🖄 🗡 🏨 🏨 17 solcap_snp. 6172. 0 0 0.8003 0.304 0.196 0.324 0.176 AA 0.7503 0.0261 1.1294 AB 0.7503 0.4769 1.5459 AA 0.7503 0.0 0.5 18 solcap_snp. 6468. 0 0 0.5474 0.363 0.216 0.196 0.225 BB 0.3268 1.0000 0.9459 AB 0.3268 0.6001 1.0532 AB 0.3268 10 enen 0.0 0.079 0.946 0.491 0 5051 0 0020 0 5051 n 9467 0 5051 0.0010 0.246 0.0040 < > Sample ID Call Rate p05 Gr Index Gender Rows=8303 Disp=8303 Sel=1 Filter=Filter is not active. 0.0000000 Unknown 1710 2149 Unknown $\triangleleft \triangleright \mathbf{X}$ Errors Table 3 $\times 1$ 1.0000000 Unknown 2254 ※ ● ● ● ● 計 計 紙 △ ※ 止 申 止 の f× Ⅲ ▼ √ ● ● 1 ■ 1.0000000 x3 Linknown 2288 5 x4 1.0000000 Unknown 2486 6 x5 0.0000000 Unknown 2845 Parent1; Parent1 Error Child/Re Child/Re SNP Parent2 Parent2 Error Child/Rep SNP Name Unknown Rep Parent1/Rep Rep Parent2 ×8 0.0000000 1998 Index Туре p Index p GType Index GType Index ×9 1.0000000 Unknown 2150 Index GType 8 9 Unknown ×10 2308 ^ 10 ×11 Unknown 2461 v 2755 11 x12 Unknown > > Rows=24 Disp=24 Sel=1 Filter=Filter is not active. Rows=0 Disp=0 Sel=0 Filter=Filter is not active Log **4** × 🖇 Select All 🚊 Copy 📕 Save 🗙 Clear 🛛 🏢 Grid 🛛 😵 Errors 🗥 Warnings 🔇 Info 🕔 Log ~ Time Source Severity Message 10/21/2010 11:11:3... INFO Starting Calculating statistics for SNP: General 10/21/2010 11:11:4... INFO Calculating statistics for SNP: 0 - 200 General 10/21/2010 11:11:4... INFO Calculating statistics for SNP: 7376 - 7576 General 10/21/2010 11:11:4... INFO Calculating statistics for SNP: 8091 - 8291 General 10/21/2010 11:11:4... INFO Calculating statistics for SNP: 8303 - 8303 General 10/21/2010 11:11:4... INFO Completed Calculating statistics for SNP: General Updated SNP statistics 10/21/2010 11:11:4... LOG Genera ast = 11:11 AM Frrors = 0Warn = 0Info = 39

Calling SNPs with Infinium 8303 Potato Array

SolCAP Custom potato calling file

- Based on potato diversity panel, two 4x populations and one 2x population
- http://solcap.msu.edu
- 3 Cluster Calling
 - Good 7412 (89.3%)
 - Questionable 296 (3.6%)
 - Segregation 254 (3.1%)
 - Bad 341 (4.1%)
- Call Rate for only good markers (7412)
 - >90% 7036 (94.9%)
 - >80% 228 (3.1%)
 - >70% 93 (1.3%)
 - <70% 55 (0.7%)



Scoring Diploid Potato on Infinium Array



Black = Parents (AA x BB)Red = Population (all AB) Black = Parents (AB \overline{x} BB) Red = Population (1:1 AB:BB)

Tetraploid SNP analysis



Tetraploid Segregation

		F1 Progeny								
Parent1	Parent2	AAAA	AAAB	AABB	ABBB	BBBB				
AAAA	AAAB	1	1							
AAAA	AABB	1	4	1						
AAAA	ABBB		1	1						
AAAB	AAAB	1	2	1						
AAAB	AABB	1	5	5	1					
AAAB	ABBB		1	2	1					
AAAB	BBBB			1	1					
AABB	AABB	1	8	18	8	1				
AABB	ABBB		1	5	5	1				
ABBB	ABBB			1	2	1				
ABBB	BBBB				1	1				
BBBB	AAAB			1	1					
BBBB	AABB			1	4	1				
BBBB	ABBB				1	1				

Illumina Genome Studio Software

Designed for use with diploid populations - Clusters are called as AA, AB, BB Potato is tetraploid with 5 marker classes - AAAA, AAAB, AABB, ABBB, BBBB – Nulliplex, Simplex, Duplex, Triplex, Quadriplex Codominant and dosage sensitive markers are ideal



Calling SNPs with 8303 Infinium Array

5 cluster custom calling using theta values

- Based on potato diversity panel, two 4x populations and one 2x population (same as 3 cluster calling)
- Summary of SNPs categories:
 - Total: 5031
 - 5 clusters: 2645
 - 4 clusters: 858
 - 3 clusters: 945
 - 2 clusters: 583
 - 1 cluster or bad SNPs: 3272



Ideal Marker



Diploid mapping population Yellow Samples = Parents Red Samples = Population Tetraploid mapping population Yellow Samples = Parents Red Samples = Population

This marker is ideal of both 2x and 4x germplasm. The 2x AA cluster overlaps the 4x AAAA cluster, the 2x AB cluster overlaps the 4x AABB cluster, and the 2x BB cluster overlaps the 4x BBBB cluster.

Scoring Tetraploid Potato Five cluster calling



Scoring Tetraploid Potato on Infinium Array





Double Reduction Example Tetraploid Potato on Infinium Array





Double Reduction



FIG. 52. Scheme for determining the theoretical maximum frequency of double reduction at a single locus in an autotetraploid. One of the three equally probable possible arrangements of the four alleles at one locus, designated a, b, c, and d, is shown together with crossovers between the locus and the centromeres, numbered 1 to 4. By assuming equal frequencies of alternate and two adjacent segregations of the centromeres, (1 + 4)/(2 + 3), (1 + 3)/(2 + 4), (1 + 2)/(3 + 4), respectively, for each of the three arrangements, the kinds of chromatids and their frequencies may be determined, as described in the text.

C.R. Burnham, Discussions in Cytogenetics, 1962



Double Reduction in Tetraploids

Autotetraploids can undergo double reduction that results in (the segments of) two sister chromatids being recovered in a single gamete.

For this to occur, multivalent pairing must take place with a cross-over between a locus and its centromere followed by the two pairs of chromatids passing to the same pole in anaphase I (adjacent segregation).



Distribution of Simplex SNPs with Double Reduction in PRRG

	No. of SNPs									
No. of DR	PR	RG	Total							
0	373	168	541							
1	47	68	115							
2	19	37	56							
3	32	14	46							
4	7	13	20							
5	7	8	15							
6	2	4	6							
7	0	1	1							



Double Reduction by Pseudomolecule Chromosome Position



Double Reduction

Tetraploid SNP Mapping Summary

PRRG	Remarks
No. of SNPs	
8303	Began with 8303 SNP from the Infinium SNP array
7666	Removed Questionable and Bad SNPs
7017	Removed SNPs unanchored to Pseudomolecule
6931	Removed SNPs mapped to >2 locations on PM
4604	Removed bad SNPs from custom 5 cluster calling
4212	Removed for 10% (19) or more No-calls in progeny
4168	Removed No-calls in Parents
3298	Removed homozygous by homozygous SNPs
3298	Segregating SNPs



Tetraploid Segregation in PR x RG

		RG								
		AAAA	AAAB	AABB	ABBB	BBBB	NC			
	AAAA	349	133	75	5	0	_ 1			
	AAAB	199	271	206	86	8	5			
PR	AABB	88	295	329	285	126	7			
	ABBB	11	79	232	306	288	1			
	BBBB	0	14	82	180	521	9			
	NC	2	2	7	6	2	2			

Nulliplex X Simplex800Nulliplex X Duplex371Nulliplex X Triplex38



Distribution of 800 Simplex SNPs in PRRG by Chromosome



Reference within PM_chr

Distribution of 800 Simplex SNPs in PRRG by Chromosome

		PR (M	1b posi	tion)	RG (Mb position)					
PM chr	N	Min	Max	Range	Ν	Min	Max	Range		
chr01	46	3.8	81.3	77.6	33	2.5	76.3	73.9		
chr02	56	1.5	46.4	45.0	20	1.5	42.3	40.8		
chr03	48	0.0	43.2	43.1	24	16.6	45.8	29.2		
chr04	43	2.1	63.5	61.4	50	0.0	64.2	64.1		
chr05	33	0.1	44.8	44.7	14	1.4	42.3	40.9		
chr06	47	1.6	54.9	53.2	31	0.3	54.9	54.6		
chr07	36	3.4	53.0	49.6	15	0.1	53.3	53.2		
chr08	46	0.3	43.3	43.1	12	1.7	40.8	39.1		
chr09	39	3.2	53.3	50.1	43	8.9	49.2	40.3		
chr10	43	0.7	50.9	50.2	19	0.3	50.7	50.4		
chr11	17	0.1	38.6	38.6	21	2.0	41.9	39.9		
chr12	33	3.8	53.7	49.9	31	0.2	56.4	56.2		



Single Marker ANOVA Specific Gravity 2010 Idaho

Parent	No. SNPs Total	No. SNPs P<0.05
PR	487	56
RG	313	77
Total:	800	133

- Multiple test corrections were too conservative
 - Bonferroni
 - Sidak-Dunnet





P-values for single marker ANOVA for 2010 SG in PR



PR chr01 single marker ANOVA example



 3 of 46
 SNPs in chr01 were significant at a=0.05

Simplex SNP genotype specific gravity means vs. PM chromosome position for Premier (P-value < 0.05)



 $\frac{E}{Y}$

Simplex SNP genotype specific gravity means vs. PM chromosome position for Rio Grande (P-value < 0.05)



Tetraploid Map

- http://www.bioss.ac.uk/download/tpmap
- Windows XP
- Designed for AFLP and SSR markers
- Maximum of 800 markers per project
- Maximum of 50 markers per linkage group
- Not effective for markers with double

reduction



Advantage of Sequenced Potato Genome

- Using only SNPs with known Pseudomolecule chromosome position
- Condordance evaluated in diploid population
- Physical map becomes a reference for comparison with the genetic map

PGSC0003DMB000000004	••••		- I I I I I I I I		+ + + + + + + + + + + + + + + + + + +	
PGSC0003DMB000000004 (1 + + + + + + + + + + + + + + + + + + +			120k		140k	
PGSC0003DMG400010020		PGSC0003DMG400010021 PGSC0003Dr	PGSC0003DM	6400010059	PGSC0003DMG400010022 PGSC0003DMG400010060	Ν
	110k 111k 112k			122k 123k 124k 125k 1		

(# 104k

Recoding SNPs for TetraploidMap

	А	В	С	D	E	F	G	Н	I.	J	K
1	SNP	PM_chr	PM_pos	Premier R	Rio Grand	PRRG-001	PRRG-002	PRRG-003	PRRG-004	PRRG-005	PRRG-006
2	c2_51812	chr01	3761637	AAAB	AAAA	AAAB	AAAB	AAAA		A A A B	AAAP
3	c2_45058	chr01	4714970	AAAB	AAAA	АААВ	АААБ	АААА	АААВ	AAAB	AAAB
4	c2_56842	chr01	30337216	AAAB	AAAA	AAAA	AAAB	AAAB	AAAB	AAAB	ΑΑΑΑ
5	c1_3234	chr01	74837511	AAAB	AAAA	AAAB	AAAA	AAAA	AAAA	AAAB	AAAA
6	c2_4715	chr01	77129418	AAAB	AAAA	AAAA	AAAA	AAAB		ΑΑΑΑ	AAAB
7	c2_22105	chr01	80786464	AAAB	AAAA	AAAB	AAAA	AAAB		AAAA	AAB
8	c2_51810	chr01	3761845	ABBB	BBBB	ABBB	ABBB	BBBB	ABBB	A DBB	ABBB
9	c2_55618	chr01	21514658	ABBB	BBBB	BBBB	ABBB	ABBB	ABBB	AB63	BBBB
10	c1_5477	chr01	23644208	ABBB	BBBB	BBBB	ABBB	ABBB	ABBB	ABBB	BBBB
11	c2_4665	chr01	76973071	ABBB	BBBB	BBBB	BBBB	ABBB	BBBB	BBBB	ABBB
12	c2_36495	chr01	77367370	ABBB	BBBB	ABBB	ABBB	BBBB	BBBB	ABBB	BBBB
13	c2_30955	chr01	81285208	ABBB	BBBB	ABBB	BBBB	ABBB		BBBB	ABBB
14											
15	SNP	PM_chr	PM_pos	PR	RG	PRRG-001	DRRG-002	DRRG-002	KRG-004	PRRG-005	PRRG-006
16	c2_51812	chr01	3761637	1	0	1	1	0	1	1	1
17	c2_45058	chr01	4714970	1	0	1	1	0	1	1	1
18	c2_56842	chr01	30337216	1	0	0	1	1	1	1	0
19	c1_3234	chr01	74837511	1	0	1	0	0	0	1	0
20	c2_4715	chr01	77129418	1	0	0	0	1	0	0	1
21	c2_22105	chr01	80786464	1	0	1	0	1	9	0	1
22	c2_51810	chr01	3761845	1	0	1	1	0	1	1	1
23	c2_55618	chr01	21514658	1	0	0	1	1	1	1	0
24	c1_5477	chr01	23644208	1	0	0	1	1	1	1	0
25	c2_4665	chr01	76973071	1	0	0	0	1	0	0	1
26	c2_36495	chr01	77367370	1	0	1	1	0	0	1	0
27	c2_30955	chr01	81285208	1	0	1	0	1	9	0	1

Recode simplex SNPs as dominant markers '1' or '0' for presence or absence

No-calls or missing data recoded as '9'

MICHIGAN ST

Recoding SNPs for TetraploidMap Marker data (.dat)

	A	В	С	D	E	F	G	Н	1	J	K	L	
1	184	46											
2	c2_518123	1 4											
3	1 ⁵	0	1	1	0	1	1	1	0	1	1	1	
4	c2_51810	1											
5	1	0	1	1	0	1	1	1	0	1	1	1	
6	c2_45058	1											
7	1	0	1	1	0	1	1	1	0	1	1	1	
8	c2_55618	1											
9	1	0	0	1	1	1	1	0	0	1	0	0	
10	c2_2873	1											
11	1	0	0	1	1	1	1	0	0	0	0	0	
12	c1_5477	1											
13	1	0	0	1	1	1	1	0	0	0	0	0	1
14	c2_56842	1											
15	1	0	0	1	1	1	1	0	0	0	0	0	

- Number of progeny (not including parents)
- 2. Number of markers
 - Marker name
 - Number of alleles (1)
 - Genotypes: Parent 1, Parent 2, then

progeny



TetraploidMap

🖏 TetraploidMap - PR chr01 -	46 SNPs.pro	Þj												
File Markers Analysis Help														
🐋 New Project 📔 Open Project		📄 🛃 Impo	ort Select	Markers	Cluster	Order 🛄 Map 📘	ANOVA	QTL						
Datasets (1)	Marker D	etaile (40/40 e	selected)											
PRRG PR chr01 - 46 SNPs.dat	#	SC Group	Name	Selected	Туре	# of Alleles	Phenotype 1	Phenotype 2	Genotype	Ratio	Ratio_Sig	Alpha	DR_Sig	
46 markers		1 B_B_D	c2_51812		AFLP		1 1000	0000	1000 0000	1:1	0.5967	0.0800	0.5987	~
Cluster Analysis 1	10	2 B_B_D	c2_51810		AFLP		1 1000	0000	1000 0000	1:1	0.5071	0.1000	0.5058	
Depdrograms		3 B_B_D	c2_45058		AFLP	22	1 1000	0000	1000 0000	1:1	0.5071	0.1000	0.5058	
E-Group 1 (46)	2	4 H_F_E	c2_55618		AFLP	13	1 1000	0000	1000 0000	1:1	0.0080	0.0	1.0	1
46 markers	5	SH_F_E	c2_2873		AFLP		1 1000	0000	1000 0000	1:1	0.0122	0.0	1.0	
Ordered Analysis 1	-	6H_F_E	c1_54//		AFLP		1 1000	0000	1000 0000	1:1	0.0122	0.0	1.0	
🗄 📄 Ordered Analysis 2	-		C2_30042				1 1000	0000	1000 0000	1:1	0.0051	0.0	1.0	
🖨 Ordered Analysis 3			c2_000				1 1000	0000	1000 0000	1.1	0.0030	0.0	1.0	
Summary	1	OHEE	c2_2653		AFLP		1 1000	0000	1000 0000	1:1	0.0073	0.0	1.0	
Details	1	1 H F F	c1_13430		AFLP		1 1000	0000	1000 0000	1:1	0.0030	0.0	1.0	
- 🕂 QTL Analysis 1	1	2GDB	c2 50903		AFLP		1 1000	0000	1000 0000	1:1	0,0097	0,1600	0.0352	
	1	3HFE	c1 12063		AFLP		1 1000	0000	1000 0000	1:1	0.8231	0.0	1.0	
Linkage Map 2	1	4 D_E_C	c1_4695		AFLP		1 1000	0000	1000 0000	1:1	0.1846	0.0	1.0	
Anova 1	1	5 H_F_E	c2_14493		AFLP		1 1000	0000	1000 0000	1:1	0.7642	0.0	1.0	
	1	.6 D_E_C	c1_4745		AFLP	2	1 1000	0000	1000 0000	1:1	0.2384	0.0	1.0	
	1	7 D_E_C	c2_14608		AFLP	22	1 1000	0000	1000 0000	1:1	0.2384	0.0	1.0	~
	Phenotype 10000000 Test for MLE of al DR signif Marker Ba 1 0 0 0 0 0 0 0 Posterior	Phenot O0000 double red pha and LR icance: 0. and Pattern 0 0 0 0 6 0 0 0 0 9 probabili	ype 2 1000 luction L test: 0.08 59874030706! us 35 0.4 32 0.5 ties in pres	0.2769 51294 802 198 sence of do	uble reduct	:ion:								
	Parental	genotypes	e Pro	bability	d.f.	chisquare	sig							
	1000 000	0	1.0	000000000	1.0	0.0	1.0							
	1100 000	00	0.0	000000000	1.0	120.94	0.0							
	1110 000	0	0.0	000000000	1.0	2255.61	0.0							
	And in ab	sence of d	louble reduct	ion:										
	1000 000	0	1.0	000000000	1.0	0.28	0.596701215	9685086						
	1100 000	0	0.0	000000000	1.0	158.9	0.0							

 Cluster analysis
 Marker ordering
 QTL Analysis



Linkage Analysis and QTL Mapping in Tetraploids Webinar

- Dr. Christine Hackett
 - Biomathematics and Statistics Scotland
- SolCAP workshop at the Potato Association of America meeting in August 2010
- This workshop is in two parts: linkage analysis, and QTL mapping.
- http://www.extension.org/pages/32471/linkageanalysis-and-qtl-mapping-in-tetraploids-webinar



PR chr01 46 SNPs



RSI

Comparison of 4x and 2x Populations to the Pseudomolecule



Coding Trait Data for QTL Analysis Phenotypic data (.qua)

	A	В	С	D	E	F
1	21	1 3				
2	2	vine_maturity	vine_matu	growth_h	total_yiel	SG_harves
3	Z ₁	4	2	7	9.57212	1.0895
4	2	3	1.5	6	6.59428	1.0805
5	3	8	7	5	13.5	1.099
6	4	5	4.5	7	19.9	1.0965
7	5	4	3.5	5	8.78552	1.099
8	6	7.5	7.5	4	4.8216	1.0865
9	7	5	4	5	19.40556	1.0845
10	8	4.5	4.5	4	9.91852	1.085
11	9	6.5	5.5	5	9.96232	1.0805
12	10	4	4	5	7.52976	1.096
13	11	7	6	3	9.02664	1.097
14	12	5	2	5	3.12984	1.092
15	13	4.5	2.5	7	13.7226	1.089
16	14	8	6	3	7.57512	1.091
17	15	6.5	5	4	8.90456	1.091
18	16	3	2	6	12.14684	1.079
19	17	6	4.5	5	12.31436	1.097
20	18	4	2.5	5	15.56384	1.096
21	19	5.5	4	5	6.68396	1.0815
22	20	-99	-99	-99	-99	-99
23	21	6	4	6	12.98756	1.0865

- 1. Number of traits
- 2. Progeny number
- 3. Trait name
- 4. Missing data coded as -99.0



Coupling and Repulsion Analysis

😳 TetraploidMap - PR chr01 - 46 SNPs.proj

🛃 start

File Markers Analysis Help											
📸 New Project 📔 Open Project		📄 🚵 Import 📰 S	ielect Mark	ers 🛛 📑 Cluste	er 🔡 O	rder 🔣 M	lap	ANOVA 🔛 QTL			
Datasets (1)	Ordere	i Results									
📄 PRRG PR chr01 - 46 SNPs.dat	twopoint.p	wd twopoint.out									
46 markers	Durdaward M.						-]
🖃 🧖 Cluster Analysis 1	Ordered Ma	arkers:				٦Ē	pcores:				
Summary	SC Group	Marker Name	Ratio	Genotype	P1	P2		Graph Code	Marker Name	Recom Freq	LOD Score
	B_B_D	c2_51812	1:1	1000 0000	1000	0000		000	c2_51812	0.2877	0.77
Group I (46)	B_B_D	c2_51810	1:1	1000 0000	1000	0000		001	c2_51810	0.2947	0.74
Ordered Applycic 1		CZ_40000	1:1	1000 0000	0100	0000		002	CZ_45058	0.2947	0.74
		c2_50903	1:1	1000 0000	0100	0000		011	1 5477	0.4423	0.06
		c2_2873	1:1	1000 0000	0010	0000		004	C1_0477	0.0	55 20
		c2_686	1.1	1000 0000	0010	0000		007	c2_2073	0.0055	52.60
Details		c2_000	1.1	1000 0000	0010	0000		007	c2_000	0.0055	52.07
OTL Analysis 1	HEE	c2_56842	1.1	1000 0000	0010	0000		005	c2_56842	0.0000	50.6
III Linkage Map 1	HEE	c2_52477	1:1	1000 0000	0010	0000		008	c2_52477	0.0389	41.33
III Linkage Map 2	HEE	c2_2653	1:1	1000 0000	0010	0000		009	c2_2653	0.0761	33.89
Anova 1	HFE	c1 13430	1:1	1000 0000	0010	0000		010	c1 13430	0.0879	31.26
	DEC	c1 4695	1:1	1000 0000	0001	0000		013	c1 4695	4.0E-4	7,93
	DEC	c2 14608	1:1	1000 0000	0001	0000		016	c2 14608	6.0E-4	7.03
	D_E_C	 c1_4745	1:1	1000 0000	0001	0000	-11	015	c1_4745	6.0E-4	7.03
	A_B_D	c2_37575	1:1	1000 0000	1000	0000		018	c2_37575	0.4507	0.04
	A_B_D	c1_5145	1:1	1000 0000	1000	0000		019	c1_5145	0.4507	0.04
	D_E_C	c1_11769	1:1	1000 0000	0001	0000		021	c1_11769	0.0454	3.76
	H_F_E	c1_12063	1:1	1000 0000	0010	0000		012	c1_12063	0.2241	12.81
	D_E_C	c1_5656	1:1	1000 0000	0001	0000		022	c1_5656	0.0612	3.49
	H_F_E	c2_14493	1:1	1000 0000	0010	0000		014	c2_14493	0.2554	9.98
	H_F_E	c1_4757	1:1	1000 0000	0010	0000		017	c1_4757	0.2717	8.64
	H_F_E	c2_49476	1:1	1000 0000	0010	0000		020	c2_49476	0.3478	3.76
	F_C_A	c2_2354	1:1	1000 0000	0100	0000		023	c2_2354	0.4917	0.01
	F_C_A	c1_2531	1:1	1000 0000	0100	0000	, l'	033	c1_2531	0.4862	0.03
	ECA.	c1 15872	1.1	1000.0000	0100		<u></u>	026	c1 15872	0 4862	
	c1_5477	1000 0000									
	c2_2873	1000 0000									
	0.5					_					
	-		100					50.0			
	<u> </u>				****						
	E E							<u>ම</u> 40.0 -			
	ita 0.31				11111		617				
	in an										
	E 0.2							<u></u> 20.0 -			
	e 0.1 -			-				10.0			
	LE .							10.0			
	0.0							0.0			
				Marker						Marker	

PR chr01 46 SNPs



HIGAN STATE

QTL Analysis in TetraploidMap PR chr01 46 SNPs



Summary

- Showed five cluster calling of SNPs in Genome Studio
- Showed steps towards filtering SNPs for tetraploid populations
- Demonstrated generating a map of chr01 in PR using TetraploidMap software with simplex markers
- Initiated QTL analysis with simplex markers
 - Single marker ANOVA
 - TetraploidMap



Databases and Resources

Integrated, breeder-focused resources for genotypic and phenotypic analysis at SGN and MSU

- http://solcap.msu.edu
- http://solanaceae.plantbiology.msu.edu
- http://solgenomics.net







Breeder's Toolbox

00		Sol Genomics Network				2
Image: the second se	/index.pl			Ċ	; Q- Google	*
Siller	sol genomics netv	vork	m contact help		Â	
認識の	search maps	genomes tools	S	sol search		
		SCN Breeders Toolbox		log in new user		
		Son Diceders Toolbox				
The purp progress bree <mark>d</mark> er-	ose of this page is to give breede and your feedback or suggestion friendly resource.	ers direct links to breeder-releve is are welcome to build this into	vant tools and data on SGN. to a comprehensive, easy to	. It is a work in o use and		
We love	feedback! Please contact Joyce v	an Eck with suggestions for	r improvements!			
		Traits Use of the second secon	Markers	S CT233 CT233 CD15 .Rt4gL5790 .Rt2g39780		

2012 Activities

- SNP genotyping on panels and populations is completed
 - Databases for genotypic and phenotypic data
 - Mining SNP genotype and phenotype data
 - QTL analyses
- Hands on workshops for breeders
- eXtension.org



Acknowledgments

Collaborators, OSU

David Francis Sung-Chur Sim Heather Merk

Collaborators, MSU

David Douches C. Robin Buell Candice Hansey John Hamilton Kim Felcher Alicia Massa

Collaborators, Cornell

Walter De Jong Lukas Mueller Joyce van Eck Naama Menda

Collaborators, UCD Allen Van Deynze

Kevin Stoffel Alex Kozik Jeannette Martins Collaborators, Oregon State Alex Stone John McQueen Roger Leigh





Funding USDA/AFRI

United States Department of Agriculture

National Institute of Food and Agriculture This project is supported by the Agriculture and Food Research Initiative Applied Plant Genomics CAP Program of USDA's National Institute of Food and Agriculture.



Please fill out the survey evaluation! (You will be contacted via email)

Today's presentation available at: http://www.extension.org/pages/63187

> Sign up for PBG News http://pbgworks.org

Watch past webinars and sign up for future webinars!

http://www.extension.org/pages/60426