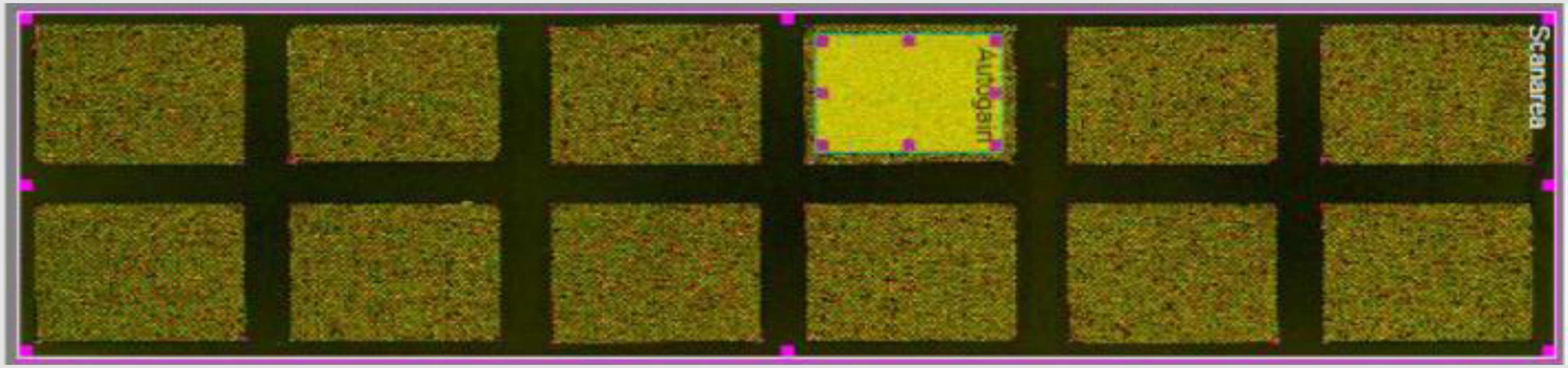
A NimbleGen CGH array for physical mapping the D-genome of hexaploid wheat with radiation hybrid lines



IWGSC workshop 2012
Thomas Drader
USDA-ARS-WRRC

Goals

- ❖Construct a physical map for the D-genome of hexaploid wheat
- High-density marker array
  - Repeat junction and gene based markers
- Wheat genetic stocks and deletion lines

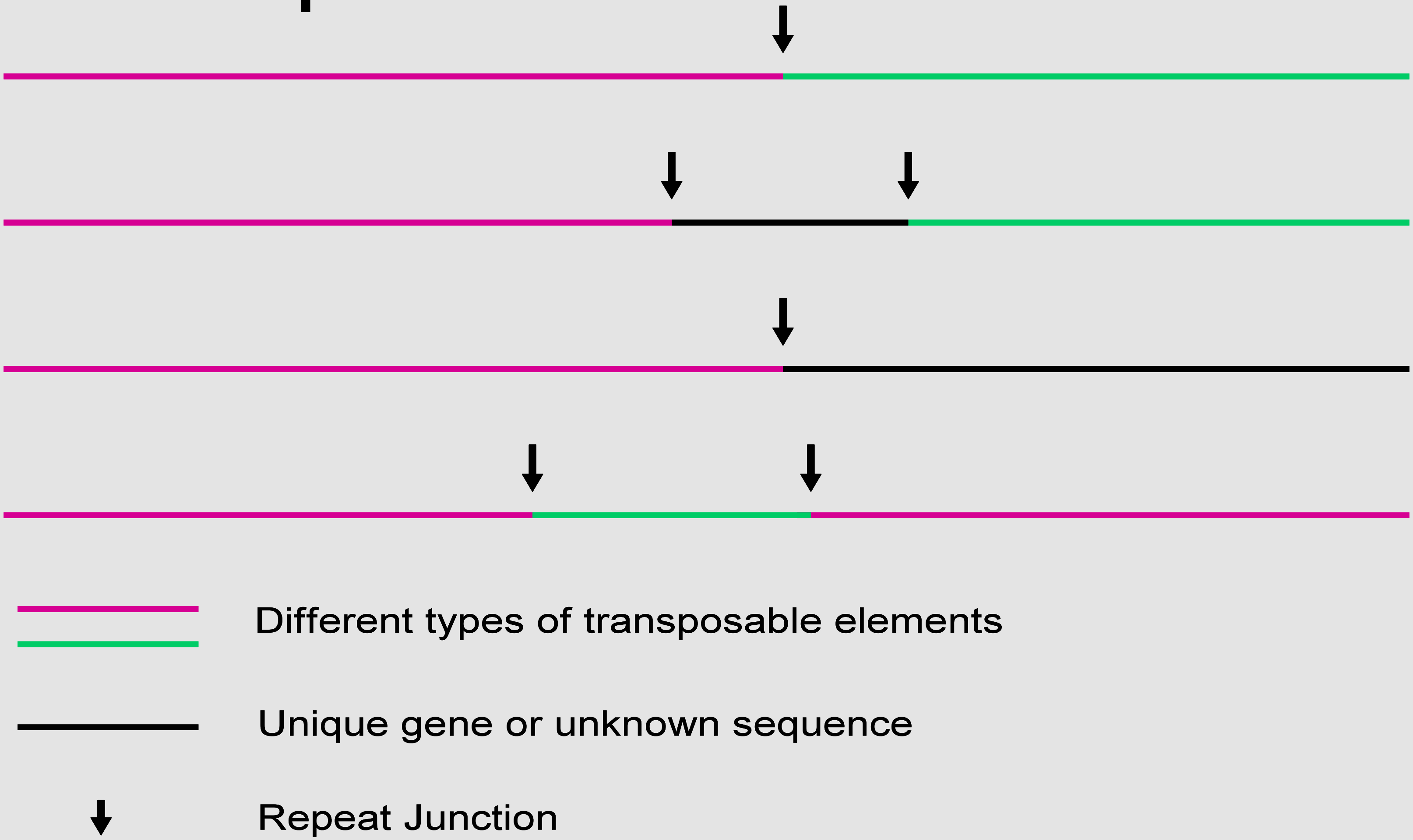
Aegilops tauschii

- D-genome progenitor
- Conserves majority of genes as the A and B genomes
- Unique transposon content (Sergeeva and Salina 2011)
- 3X 454 Genome sequence

Gene Based Markers

- Limited to Genetic space
- Not evenly distributed along the chromosome
- High potential for cross reaction between A, B, and D homeologs

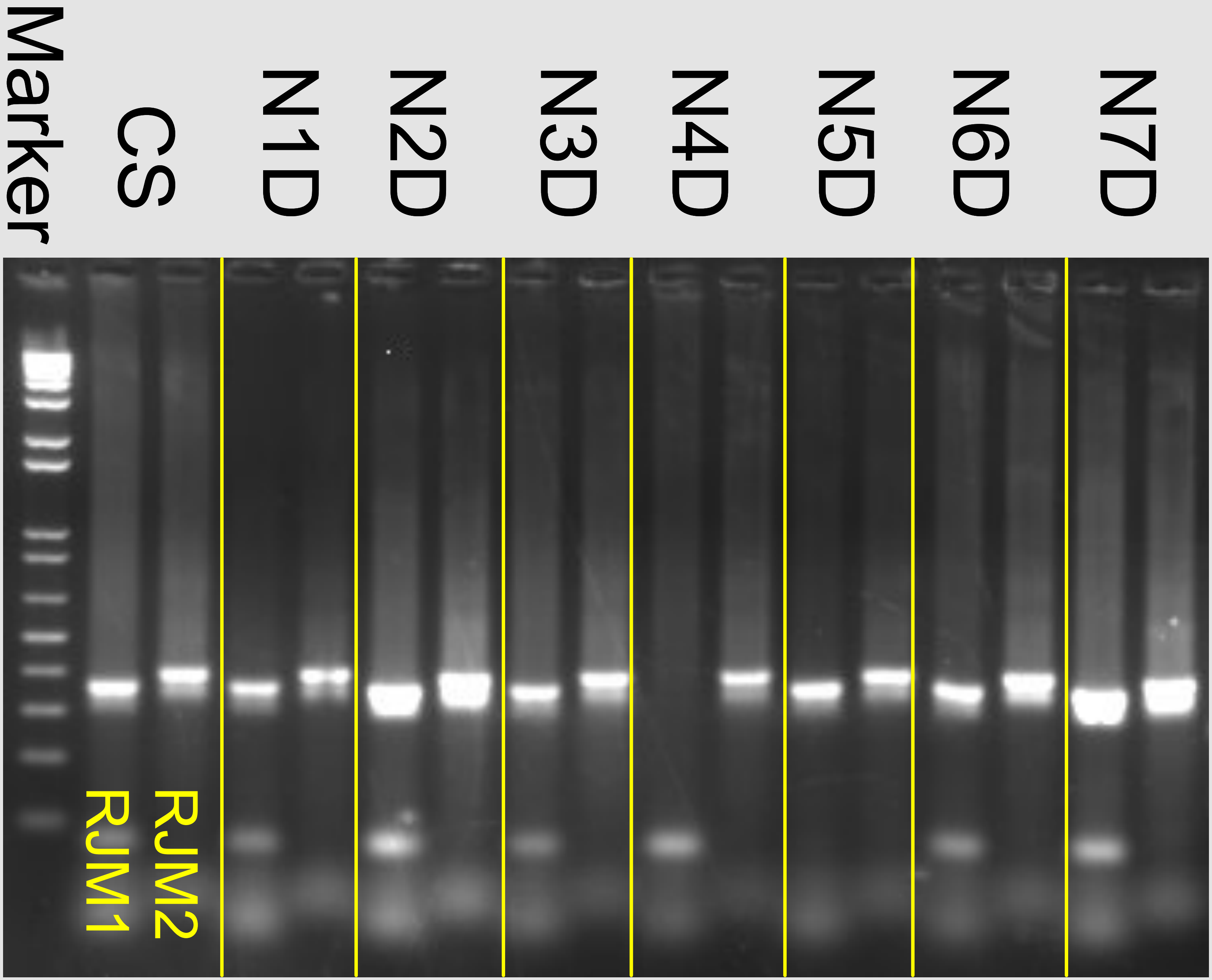
Repeat Junction Marker



Different types of transposable elements

Unique gene or unknown sequence

Repeat Junction


Marker
CS
N1D
N2D
N3D
N4D
N5D
N6D
N7D
RJM2
RJM1

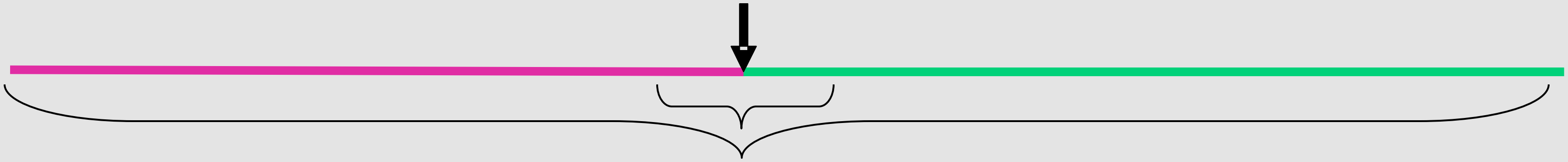
Ae. Tauschii Repeat Junction

- 3X 454 Genome sequence of *Aegilops tauschii* accession AL8/78
- Screened using RJPrimer program (http://probes.pw.usda.gov/RJPrimers/)
- High-copy junctions were identified via BLASTn and removed
- The sequences were verified using Illumina sequencing reads

NimbleGen CGH Array

- High-density microarray for variable length oligos
- Allow for the detection of copy number differences between a test and reference sample
- Available for custom array development

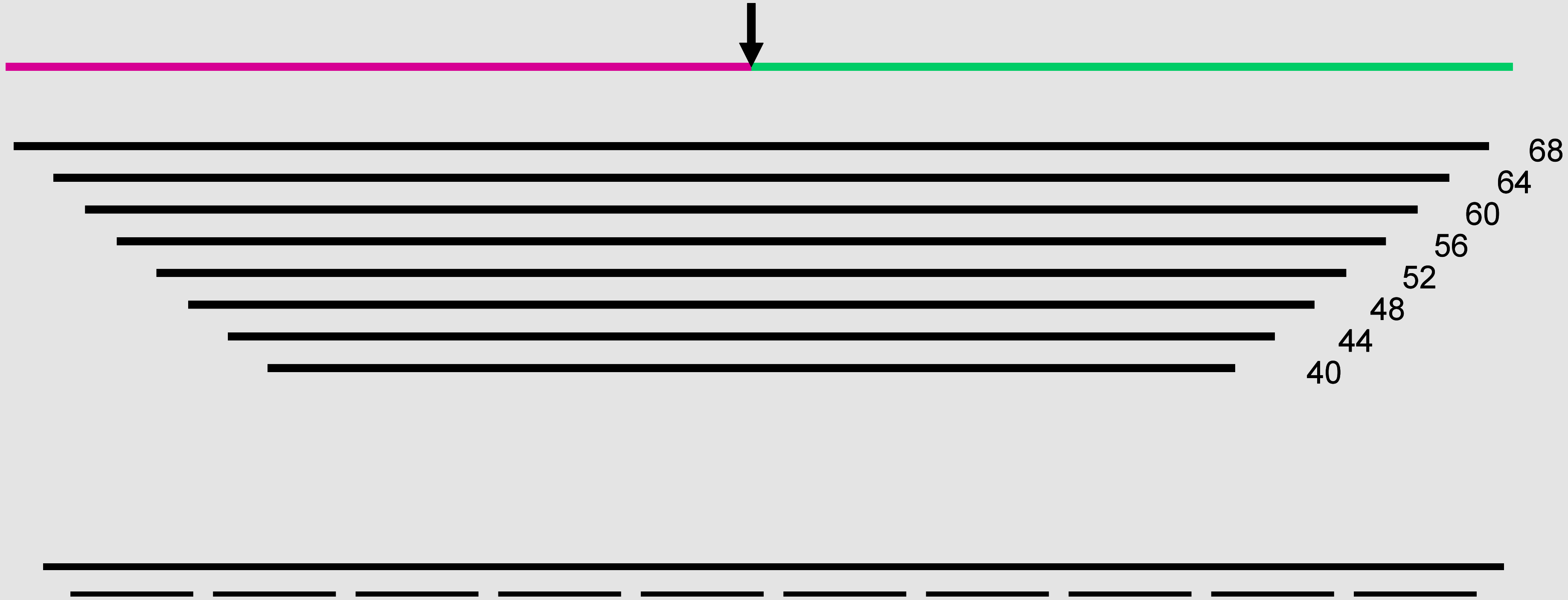
Array Design



- At what length is the oligo specific?
- Too short = non-specific hybridization
- Too long = non-specific binding to repetitive elements

- Could higher temperature positively influence the longer probe?

Test Array at 42C and 47C



Assignment Criteria

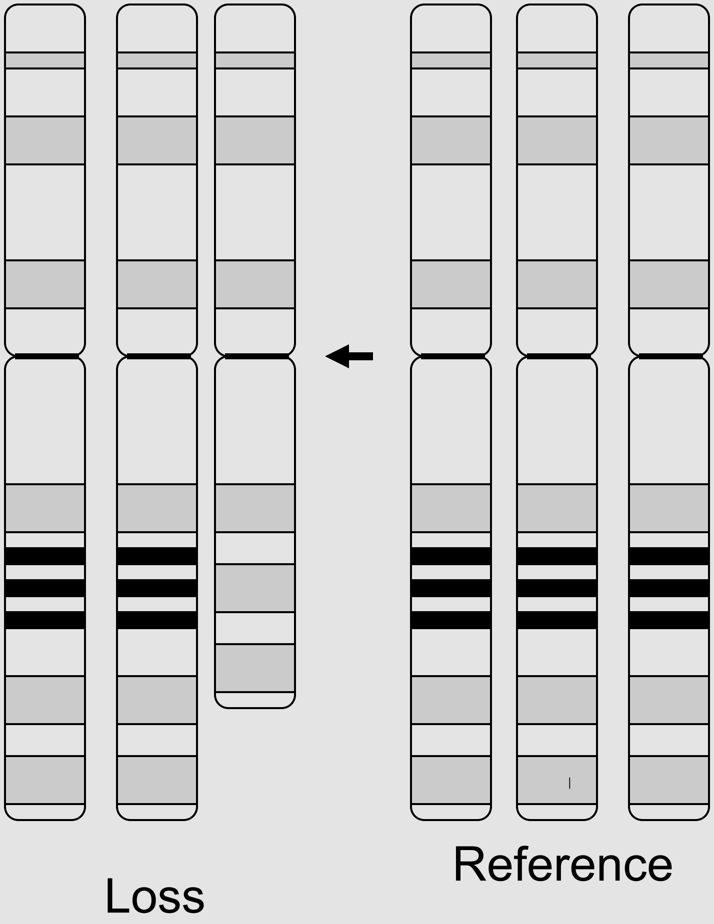
- Initially started with signal loss of 50% but cross hybridization from the A and B genome decrease the signal loss to 1/3
- Minimum Loss of signal 20% was chosen
- Student's T-test significant p-value < 0.05

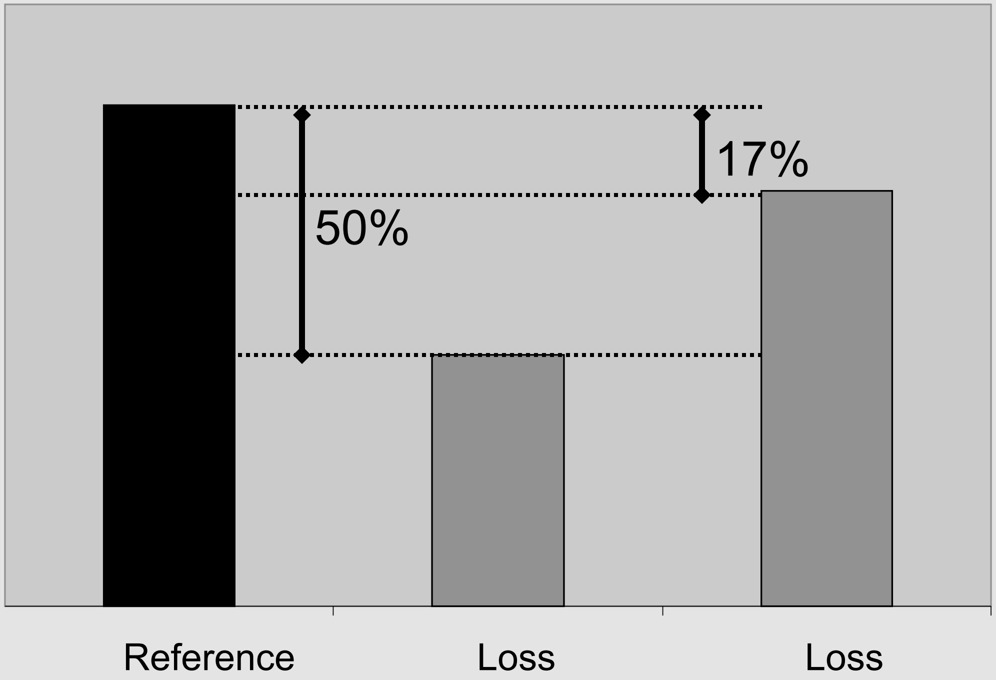
Probe Assignment

42C	Genes	_40	_44	_48	_52	_56	_60	_64	_68
N1D_532	112	97	101	97	103	105	104	132	162
N2D_532	126	132	134	124	114	117	95	82	65
N3D_532	113	131	149	138	122	110	99	85	76
N4D_532	103	120	119	114	109	106	82	72	72
N5D_532	383	123	131	118	116	123	93	91	83
N6D_532	94	203	301	329	438	382	377	328	233
N7D_532	133	178	215	212	217	223	179	148	103
%Total Assigned	76.55%	17.57%	20.53%	20.21%	21.76%	20.81%	18.37%	16.74%	14.17%

47C	Genes	_40	_44	_48	_52	_56	_60	_64	_68
N1D_532	104	12	23	43	49	66	73	100	136
N2D_532	114	10	38	77	100	115	132	130	134
N3D_532	78	14	34	68	85	102	122	117	127
N4D_532	64	55	49	74	88	95	111	109	103
N5D_532	116	12	32	60	69	97	100	106	102
N6D_532	85	22	43	65	81	119	132	156	143
N7D_532	120	30	44	69	100	104	133	128	118
%Total Assigned	48.99%	2.77%	4.69%	8.14%	10.21%	12.46%	14.32%	15.10%	15.41%

Cy3/Cy5 Ratio in Hexaploid Wheat





Cutoff for signal loss

_52	**<1.0**	**<0.9**	**<0.8**	**<0.7**	**<0.6**	**<0.5**	**<0.4**
N1D_532	187	170	103	51	16	8	1
N2D_532	166	157	114	66	28	12	3
N3D_532	197	174	122	67	23	8	3
N4D_532	151	148	108	53	21	7	1
N5D_532	161	154	116	63	24	7	2
N6D_532	520	513	438	235	74	13	2
N7D_532	252	246	217	135	54	17	4
Total	1634	1562	1218	670	240	72	16
%Assign	29.17%	27.88%	21.74%	11.96%	4.28%	1.29%	0.29%

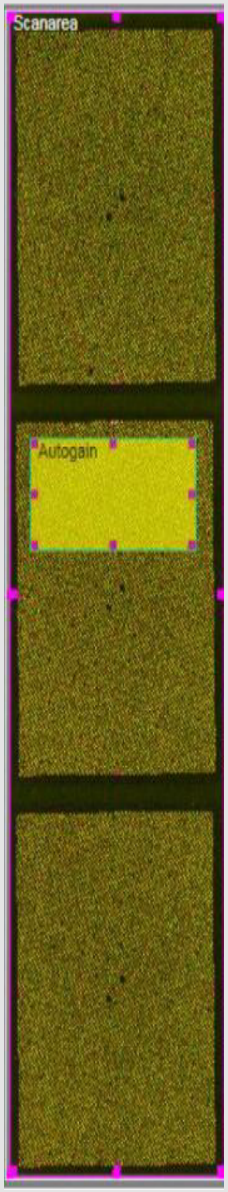
genes	**<1.0**	**<0.9**	**<0.8**	**<0.7**	**<0.6**	**<0.5**	**<0.4**
N1D_532	119	117	112	105	85	62	40
N2D_532	130	130	126	108	80	44	17
N3D_532	116	115	113	90	70	50	31
N4D_532	110	110	103	85	55	38	22
N5D_532	415	412	383	288	115	65	34
N6D_532	95	95	94	86	66	46	26
N7D_532	138	138	133	123	97	66	40
Total	1123	1117	1064	885	568	371	210
Match to SNP	801	799	779	699	521	349	196
%Assign	46.79%	46.54%	44.33%	36.88%	23.67%	15.46%	8.75%
%Match	71.33%	71.53%	73.21%	78.98%	91.73%	94.07%	93.33%

Design summary

- 42C provides greater number of assigning probes
- Oligo length of _52 provides highest number of assigning probes
- 20% signal loss provides 40% markers assigned with accuracy of 73%

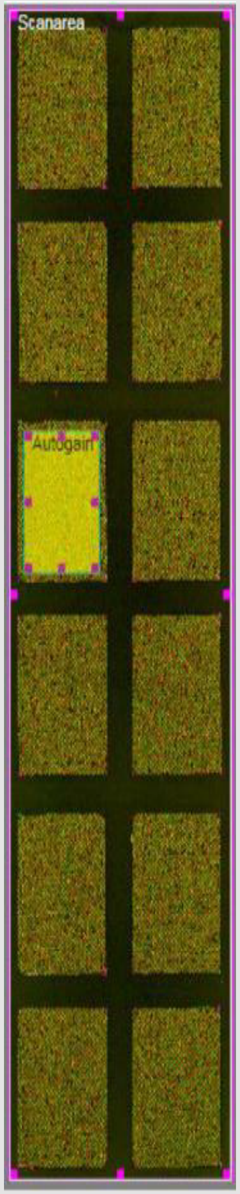
Screening Array

- 3 X (3 X ~240,000 probe) array from 3X genome coverage of *Ae. tauschii* and 6,700 gene based markers
- Screened in duplicate with control Chinese Spring and Nulli-tetrasomic lines 1D-7D
- Successful markers must assign to the same chromosome in duplicate and cannot assign to Chinese Spring



Screening Results

- 30,900 repeat junction markers
- 6,330 genes with 2-3 probes/gene
- In total 46,221 probes
- 12 X (3 X 45,000 probe) Mapping array



Mapping

NT-1D	4268	DT-1DL	1736
NT-2D	2870	DT-1DS	2215
NT-3D	2108	DT-2DL	1271
NT-4D	2629	DT-2DS	1597
NT-5D	4294	DT-3DL	654
NT-6D	3272	DT-3DS	1281
NT-7D	4450	DT-4DL	1026
Total	23891	DT-4DS	1522
		DT-5DL	2833
		DT-6DL	1288
		DT-6DS	1879
		DT-7DS	3456

Bin Mapping

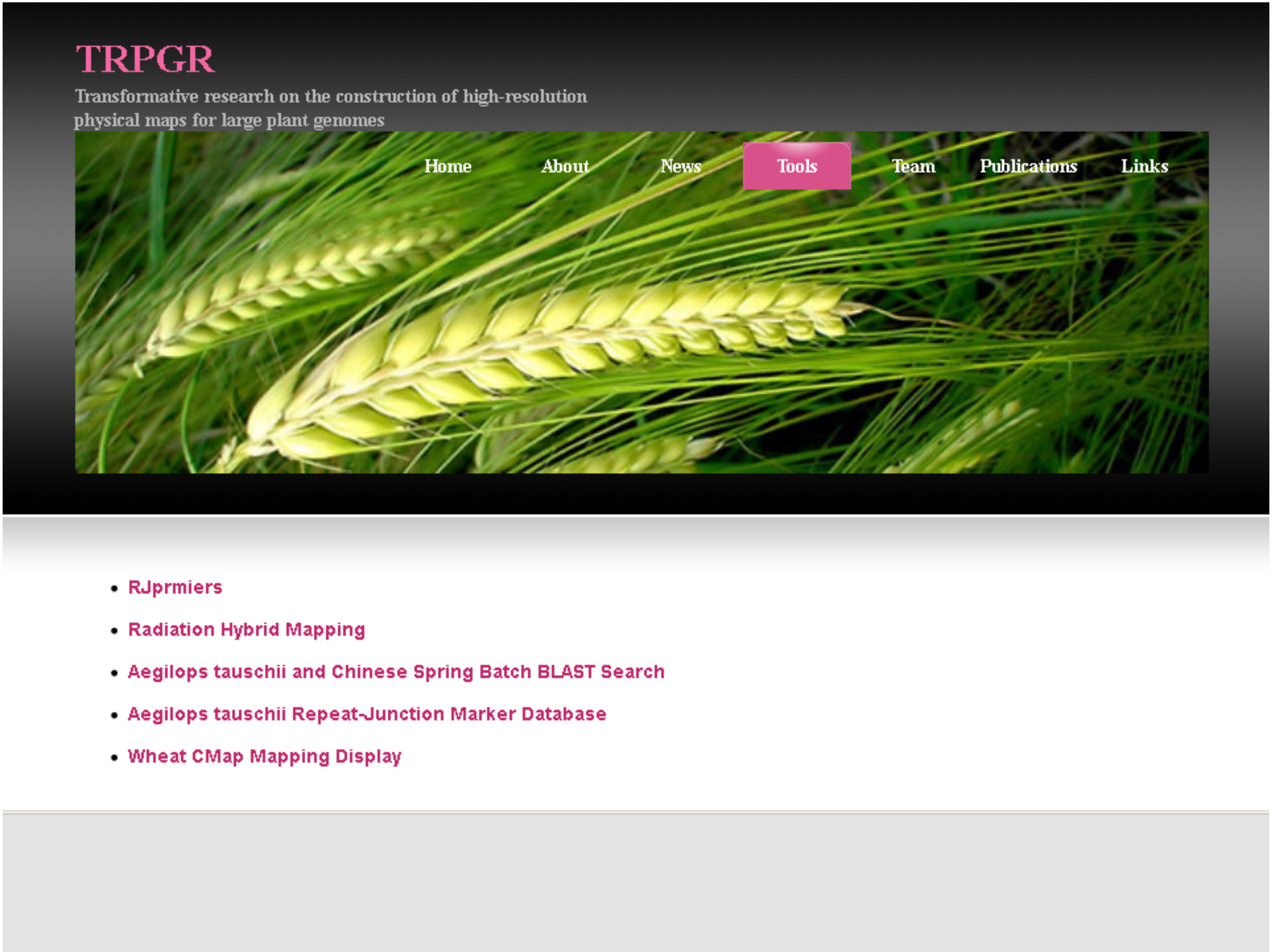
1DL-2	3398	2DL-3	3834	3DL-2	4242	4DL-12	2307	5DL-1	3746	6DL-1	2281	7DL-2	3496
1DL-4	4403	2DL-6	2817	3DL-3	3325	4DL-13	3022	5DL-5	3801	6DL-10	2119	7DL-3	3108
1DS-1	2673	2DL-9	3027	3DS-3	3688	4DL-9	3480	5DL-7	2056	6DL-11	3347	7DL-5	4452
1DS-5	2080	2DS-1	3259	3DS-6	2877	4DS-1	2899	5DL-9	3489	6DL-12	2764	7DS-4	3869
		2DS-5	3220	3DS-7	3149	4DS-2	1449	5DS-1	3455	6DL-6	3863	7DS-5	4438
						4DS-3	2248	5DS-2	2960	6DS-2	2943	7DS-6	3047
								5DS-5	3269	6DS-4	2099		
										6DS-6	2417		

Summary and Future Work

- Goal of 8,000 markers
- 23,891 markers mapped to specific bins
- ~1200 Radiation hybrid lines will be used for higher resolution physical map

- Data is available at:

avena.pw.usda.gov/RHmapping/tools.html



TRPGR

Transformative research on the construction of high-resolution physical maps for large plant genomes

- RJprmiers
- Radiation Hybrid Mapping
- Aegilops tauschii and Chinese Spring Batch BLAST Search
- Aegilops tauschii Repeat-Junction Marker Database
- Wheat CMap Mapping Display

NimbleGen Array

- Design is available through December of 2012
- After December of 2012 the array will be available in an alternate format
- Please contact: thomas.drader@ars.usda.gov

Thank you

USDA-ARS Western Regional Research Center
- Yong Gu
- Lingli Dong
- Yi Wang
- Gerard Lazo

University of California, Davis
- Ming-Cheng Luo
- Jan Dvorak

Oregon State University
- Vijay K. Tiwari
- Jeff M. Leonard

North Dakota State University
- Shahryar Kianian
- Ajay Kumar
- Muhammad Iqbal
- Anne Denton