

# ILLUMINA SNPs IN BARLEY

David Marshall  
Genetics Programme



# Barley Sequence Based Markers

- The recent availability of > 5000 sequence based markers in barley
- What can we learn about the relationship between the rice and barley genomes from our new SNP/SFP based barley map ?
- What does it tell about the composition of the barley genome ?
- How many genes are there in barley and where are they located
- How good is the coverage of existing barley maps
- What are the likely patterns of LD in barley ?

# Barley Sequence based markers

- Illumina Golden Gate OPAs

1536 SNPs per OPA

One SNP per gene on 1<sup>st</sup> 2 x Pilot OPA, multiple SNPs on Pilot OPA3

Genotyping material - three mapping populations and 102 barley varieties

- Production set of 2 x OPAs derived from initial 3-4 Pilots

- Affymetrix Barley1 GeneChip

SFPs from Steptoe x Morex mapping population

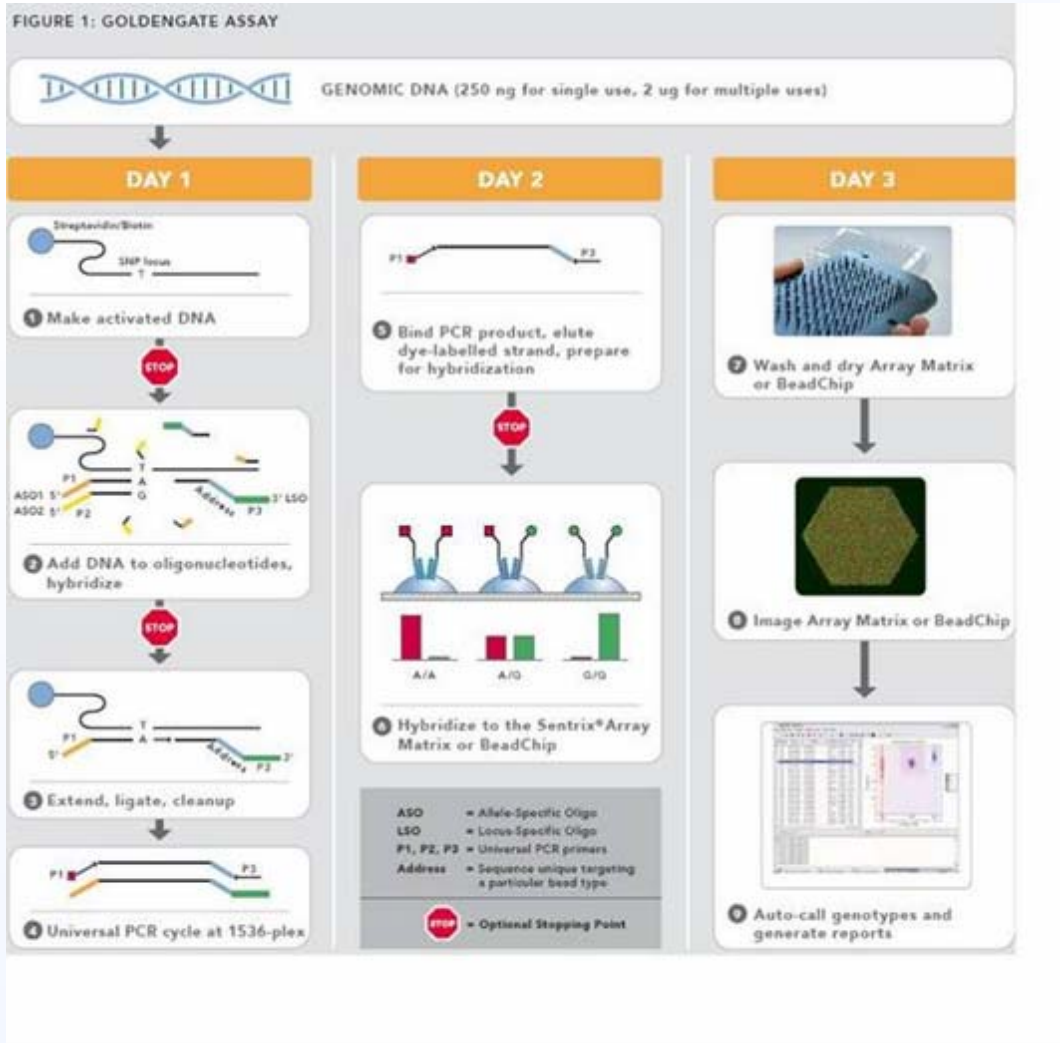
>3,000 SFPs from two different developmental stages

Majority of SFPs associated with structural gene but significant proportion do not correspond to site of probe sets

- Currently we have >5,000 sequence based markers mapped in barley mapping populations

- Large scale genotyping projects underway in the UK (AGOUEB) and US (USDA Barley CAP)

# Barley Illumina Genotyping



- SNPs design against >400,000 ESTs
- Resequencing of abiotic stress related genes
- UCR/SCRI/IPK collaboration
- 3 x 1536 Pilot Oligo Pooled Arrays (OPAs)
- 2 x 1536 SNP Production arrays.

# Illumina design

- SNPs discovery either through eSNP mining from >400,000 barley SNPs or re-sequenced in test panel of genotypes (8 or 25)
- Fasta file with SNP marked as [C/G] any other polymorphic sites identified with ambiguity codes
- Design score – minimum threshold for individual assay and maximum number of low quality SNPs to ensure success of 1536
- Individual SNP assay described in terms of 3 primers, Fasta design file this may or may not have a simple relationship to cDNA or genomic sequence (discuss)

# Barley SNP Database

**Details of Primer Pair 'ABC00086\_L01R01'**

Pair	Est Pred Size	Submitter	Date	Primer	Sequence	Start
ABC00086_L01R01	418	Nils Rostoks	2003-11-12	ABC00086_L01	CCCCCAATGATGTGCGCTTT	870
				ABC00086_R01	GGGGCTCACATGACCCCTAG	128

**Comment History of 'ABC00086\_L01R01'**

Date	Submitter	Comment
2003-11-12	Nils Rostoks	SCRI SNP discovery project in abo

**Polymorphisms for ABC00086**

Schematic of mutation locations within contig assembly  
(Hover mouse over vertical lines to highlight SNP locations in the table)

Golden PromiseL  
Golden PromiseR  
HS92R  
HS92L  
MorexR  
MorexL  
OWB DPR  
OWB DL  
OpticL  
OpticR  
OWB RPR  
OWB RL  
SteppeL  
SteppeR

**Polymorphic base positions**

Positions indicated by red arrows are mapped SNPs.

SNPs detected by primers

Consensus sequence positions	2	3	3	3	4	5	6	6	6	8
Golden Promise (L01R01) L01_1	T	A	C	C	A	C	A	T	G	T
Golden Promise (L01R01) R01_1	T	A	C	C	A	C	A	T	G	T
HS92 (L01R01) R01_1	T	A	C	C	A	C	A	T	G	T
HS92 (L01R01) L01_1	T	A	C	C	A	C	A	T	G	T
Morex (L01R01) R01_1	T	A	C	C	A	C	A	T	G	T
Morex (L01R01) L01_1	T	A	C	C	A	C	A	T	G	T
OWB D (L01R01) R01_1	T	A	C	C	A	C	A	T	G	T
OWB D (L01R01) L01_1	T	A	C	C	A	C	A	T	G	T
Optic (L01R01) L01_1	T	A	C	C	A	C	A	T	G	T
Optic (L01R01) R01_1	T	A	C	C	A	C	A	T	G	T
OWB R (L01R01) R01_1	T	A	C	C	A	C	A	T	G	T
OWB R (L01R01) L01_1	T	A	C	C	A	C	A	T	G	T

Primers Information etc

Visualisation of SNPs within re-sequenced lines

# Barley Gene Register

Show only hits with BLAST score >= 
Barley hit cutoff:  Rice hit cutoff:

Show multiple alignments (cutoff >= 50)

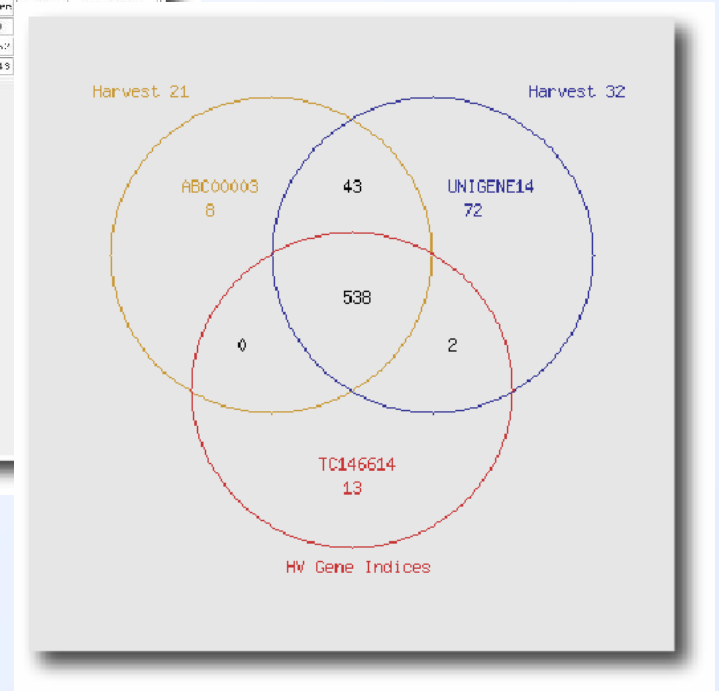
Display BLAST hits to contigs:

Available datasets (click for more info):  
[HarVEST 21](#) (eq. ABC00001) [HarVEST 32](#) (ec. UNIGENE1)  
[TIGR Hv gl](#) (eq. TC\_17035) [Rice PP1](#) (ec. Os11G05133)

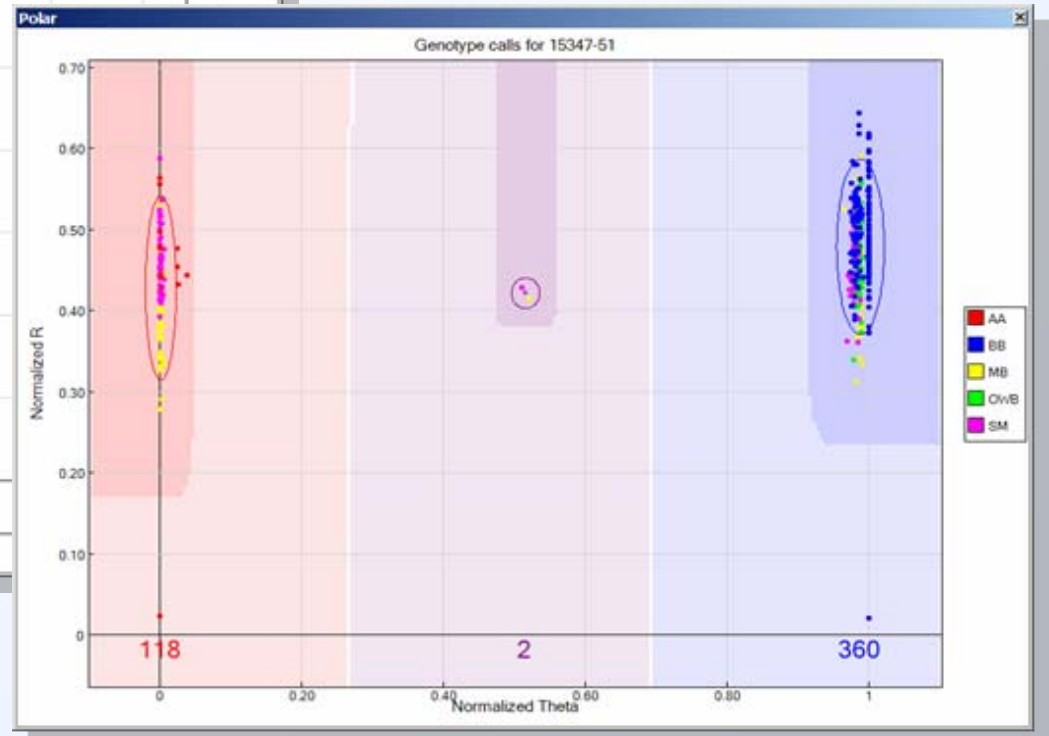
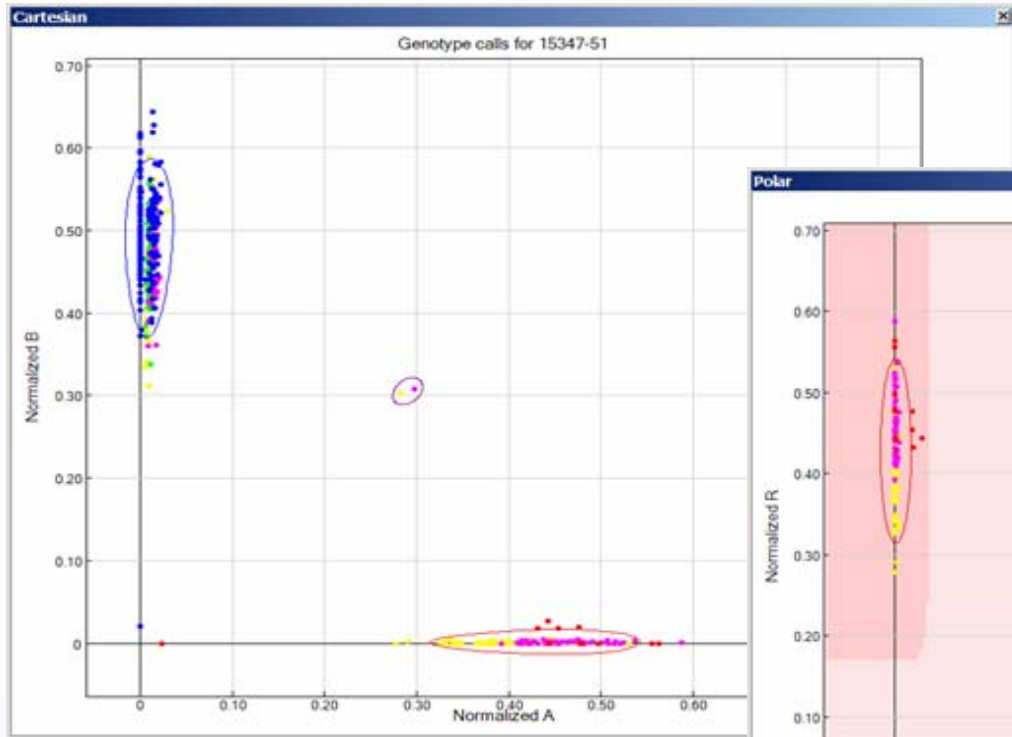
Select which databases to search:  
 HarVEST 21  HarVEST 32  TIGR Hv Gene Indices  Rice PseudoPeptides v4

### Hits related to ABC08523 \*

HarVEST 21					HarVEST 32				TIGR Hv Gene Indices				Rice PseudoPeptides v4		
Hit	Hv Affy	Acc. ID	Score	E-value	Identities	Hit	Score	E-value	Identities	Hit	Score	E-value	Identities	Hit	Score
ABC08523			0.0	14.4	11/11/111 (97%)	U9421	0.0	2319	1181/1131 (99%)	TC133122	0.0	2782	1515/1556 (97%)	Os11G05100	0.0
ABC08525			0.0	1053	555/623 (89%)	U9422	0.0	1473	815/833 (98%)	AV94846	0.0	406	252/276 (91%)	Os11G05101	0.1162
ABC08498	<a href="#">GO</a>	BC0148130	0.126	481	314/333 (94%)	U9420	0.0	1068	601/523 (56%)					Os12G05500	0.145

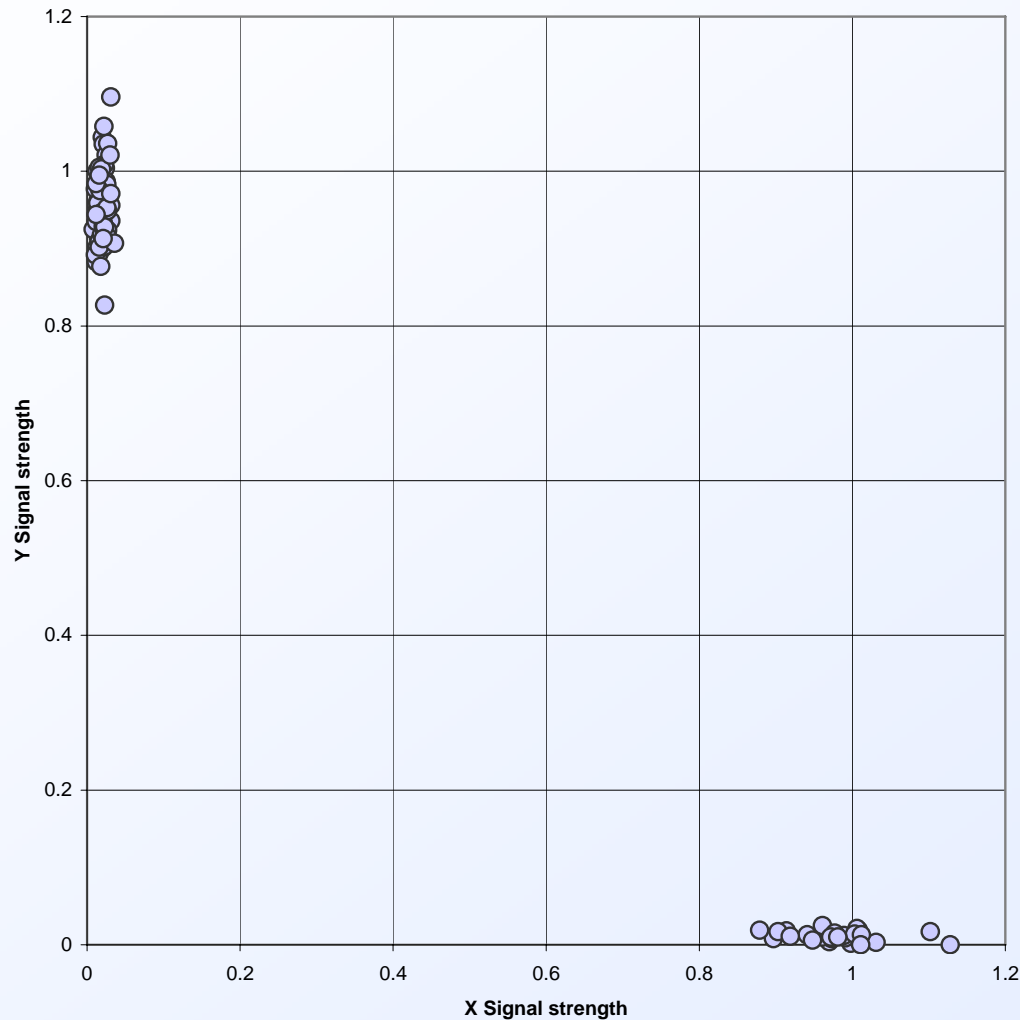


# SNP Genotype Calling



# There is nothing special about the Illumina software

Plot of Normalised X and Y signal strength for Locus 4944-2118



# ABI SNP Calling

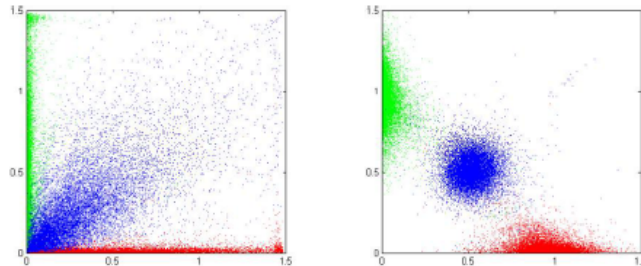


Figure 3-22 Pre- and post- normalization data. Shown are 80,000 measurements of SNPlex System data overlaid in normalized allele coordinates. In the left plot, data are normalized only by run. In the right, the data are fully normalized by sample and allele.

## Genotyping

The random error remaining in the data after normalization is relatively small when compared with the separation of the genotype coordinates, as seen in Figure 3-23. Each measurement is assigned the genotype for which it has the highest probability.

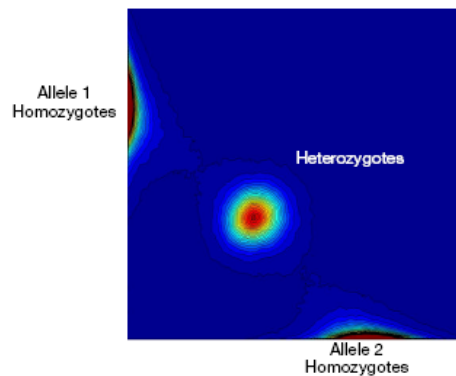


Figure 3-23 Training data distribution

# Exporting Calls

Microsoft Excel - example illumina output

File Edit View Insert Format Tools Data Window Help Adobe PDF

Type a question for help

Calibri 11

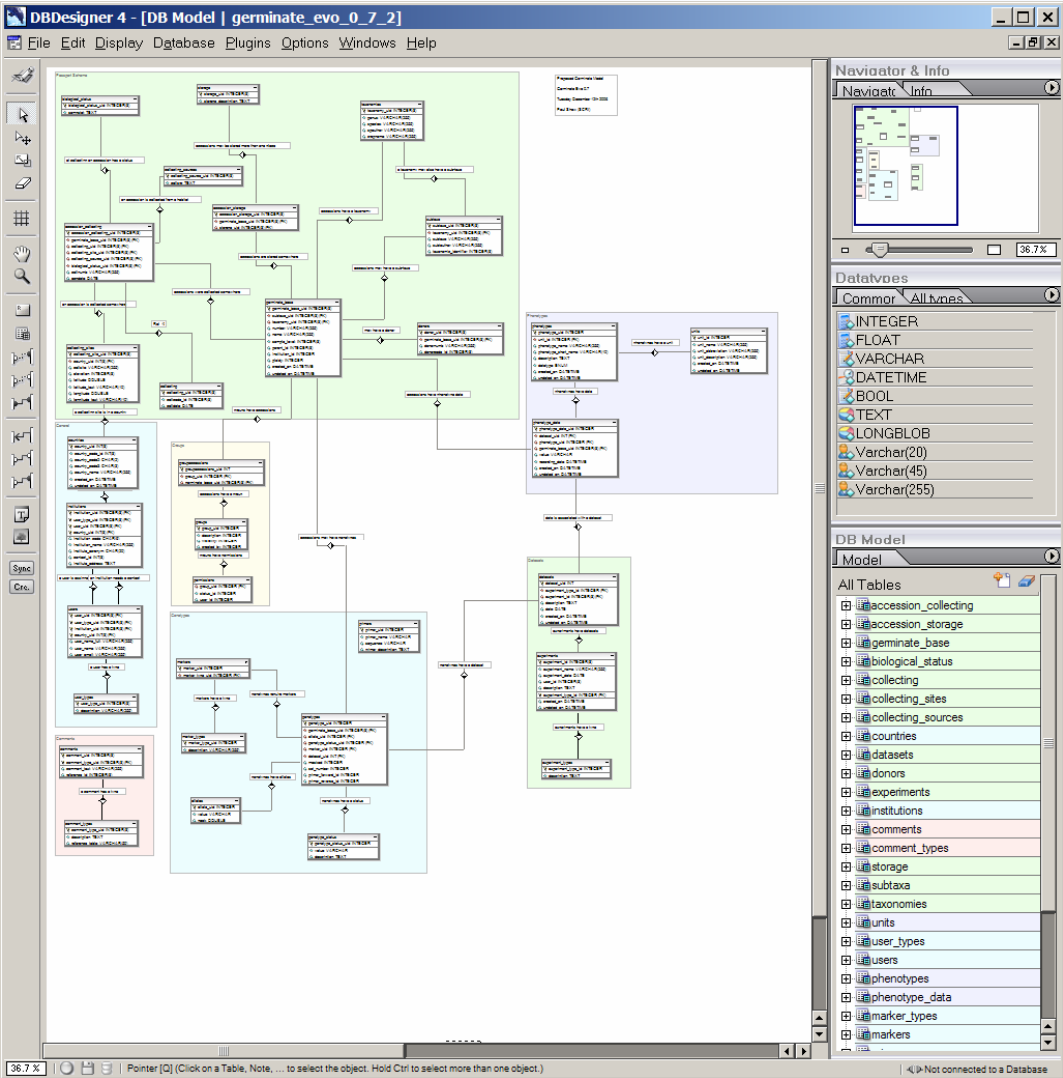
Reply with Changes... Egd Review...

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
			DNA	DNA Aux	Locus ID	Locus Name	Locus Aux	Chromosome	Coordinate	Call	Submitted Strand Nucleotide Call	Score	X	Y	Xraw	Yraw	
1																	
2	1	1338450_R001_C001	0	3	1074-992	0		0	AA	TT	0.63	0.2	0.04	2761	895		
3	2	1338450_R001_C002	0	3	1074-992	0		0	AA	TT	0.63	0.27	0.04	3534	812		
4	3	1338450_R001_C003	0	3	1074-992	0		0	AA	TT	0.63	0.31	0.05	3501	903		
5	4	1338450_R001_C004	0	3	1074-992	0		0	AA	TT	0.63	0.24	0.04	3087	788		
6	5	1338450_R001_C005	0	3	1074-992	0		0	AA	TT	0.63	0.43	0.05	5105	954		
7	6	1338450_R001_C006	0	3	1074-992	0		0	AA	TT	0.63	0.39	0.04	4195	795		
8	7	1338450_R001_C007	0	3	1074-992	0		0	AA	TT	0.63	0.44	0.08	4509	1243		
9	8	1338450_R001_C008	0	3	1074-992	0		0	AA	TT	0.63	0.42	0.05	4784	948		
10	9	1338450_R001_C009	0	3	1074-992	0		0	AA	TT	0.63	0.38	0.04	3768	667		
11	10	1338450_R001_C010	0	3	1074-992	0		0	AA	TT	0.63	0.42	0.07	4636	1124		
12	11	1338450_R001_C011	0	3	1074-992	0		0	AA	TT	0.63	0.39	0.06	4249	995		
13	12	1338450_R001_C012	0	3	1074-992	0		0	AA	TT	0.63	0.34	0.09	4205	1611		
14	13	1338450_R002_C001	0	3	1074-992	0		0	AA	TT	0.63	0.22	0.05	2871	979		
15	14	1338450_R002_C002	0	3	1074-992	0		0	AA	TT	0.63	0.31	0.05	3885	1079		
16	15	1338450_R002_C003	0	3	1074-992	0		0	AA	TT	0.63	0.28	0.04	3365	754		
17	16	1338450_R002_C004	0	3	1074-992	0		0	AA	TT	0.63	0.37	0.04	4708	929		
18	17	1338450_R002_C005	0	3	1074-992	0		0	AA	TT	0.63	0.29	0.05	3160	843		
19	18	1338450_R002_C006	0	3	1074-992	0		0	AA	TT	0.63	0.34	0.05	3903	865		
20	19	1338450_R002_C007	0	3	1074-992	0		0	AA	TT	0.63	0.39	0.05	4101	813		
21	20	1338450_R002_C008	0	3	1074-992	0		0	AA	TT	0.63	0.37	0.05	4331	1004		
22	21	1338450_R002_C009	0	3	1074-992	0		0	AA	TT	0.63	0.33	0.06	3304	936		
23	22	1338450_R002_C010	0	3	1074-992	0		0	AA	TT	0.63	0.36	0.06	3931	1117		
24	23	1338450_R002_C011	0	3	1074-992	0		0	AA	TT	0.63	0.37	0.08	3994	1312		
25	24	1338450_R002_C012	0	3	1074-992	0		0	AA	TT	0.63	0.32	0.08	3647	1446		
26	25	1338450_R003_C001	0	3	1074-992	0		0	AA	TT	0.63	0.25	0.04	3171	765		
27	26	1338450_R003_C002	0	3	1074-992	0		0	AA	TT	0.63	0.36	0.06	4304	1087		
28	27	1338450_R003_C003	0	3	1074-992	0		0	AA	TT	0.63	0.32	0.03	4532	843		
29	28	1338450_R003_C004	0	3	1074-992	0		0	AA	TT	0.63	0.34	0.04	3845	857		

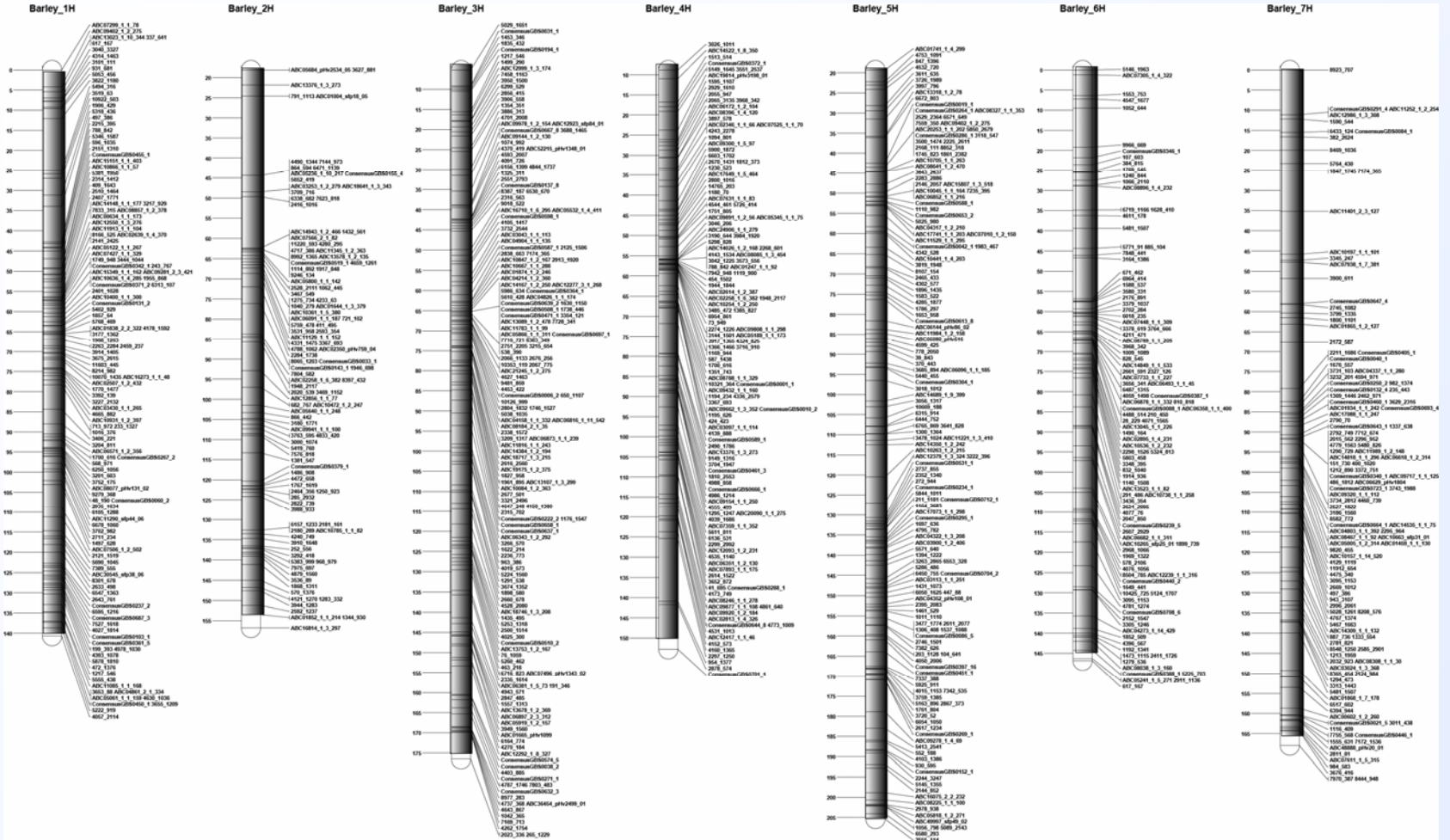
Sheet1 / Sheet2 / Sheet3

Ready

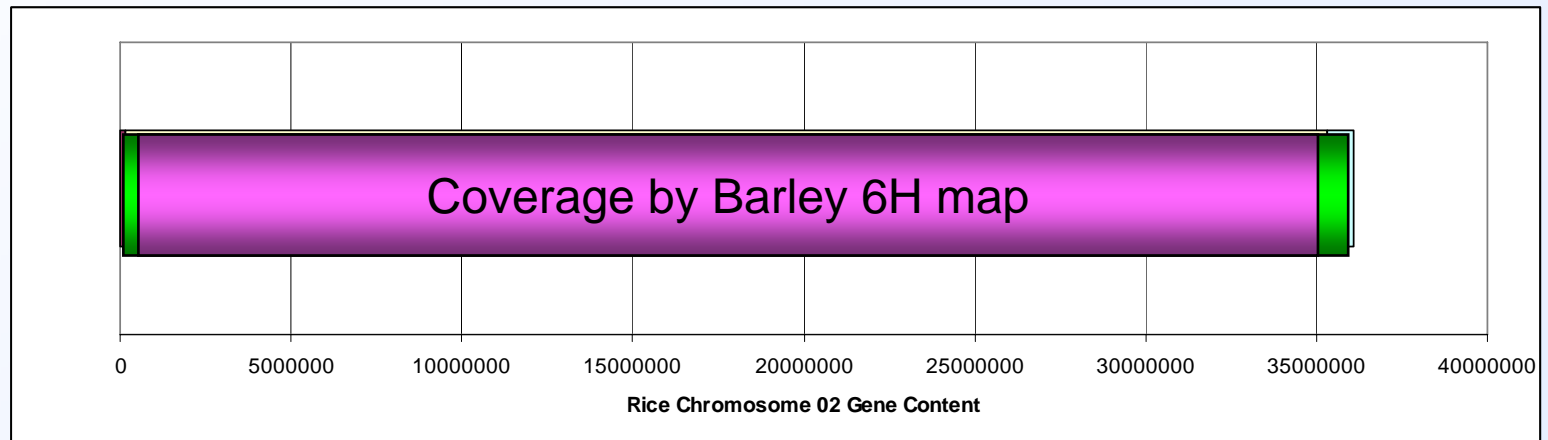
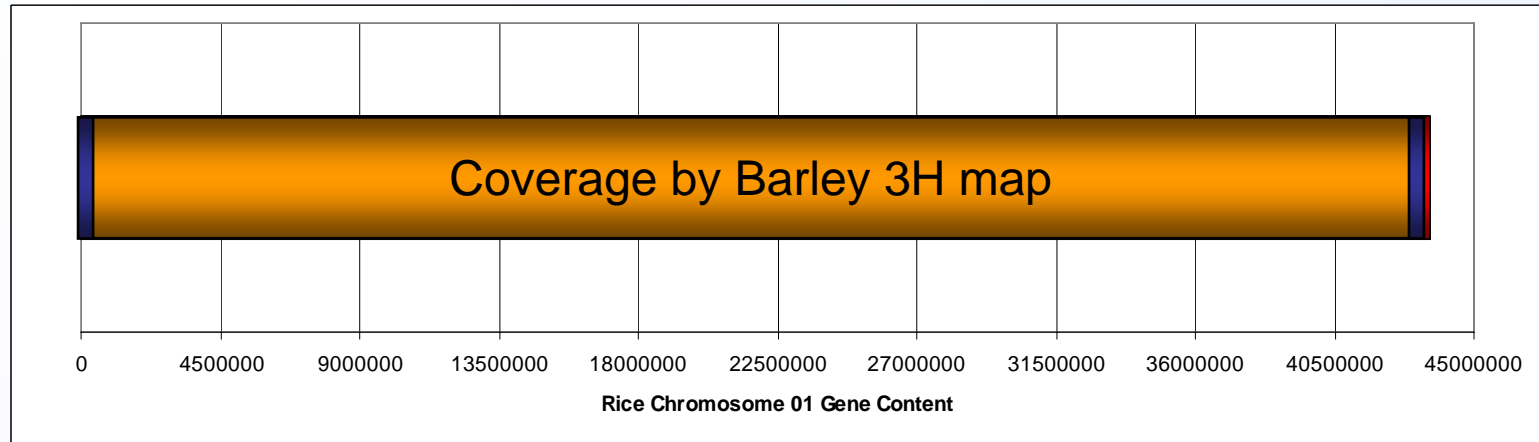
# Germinate

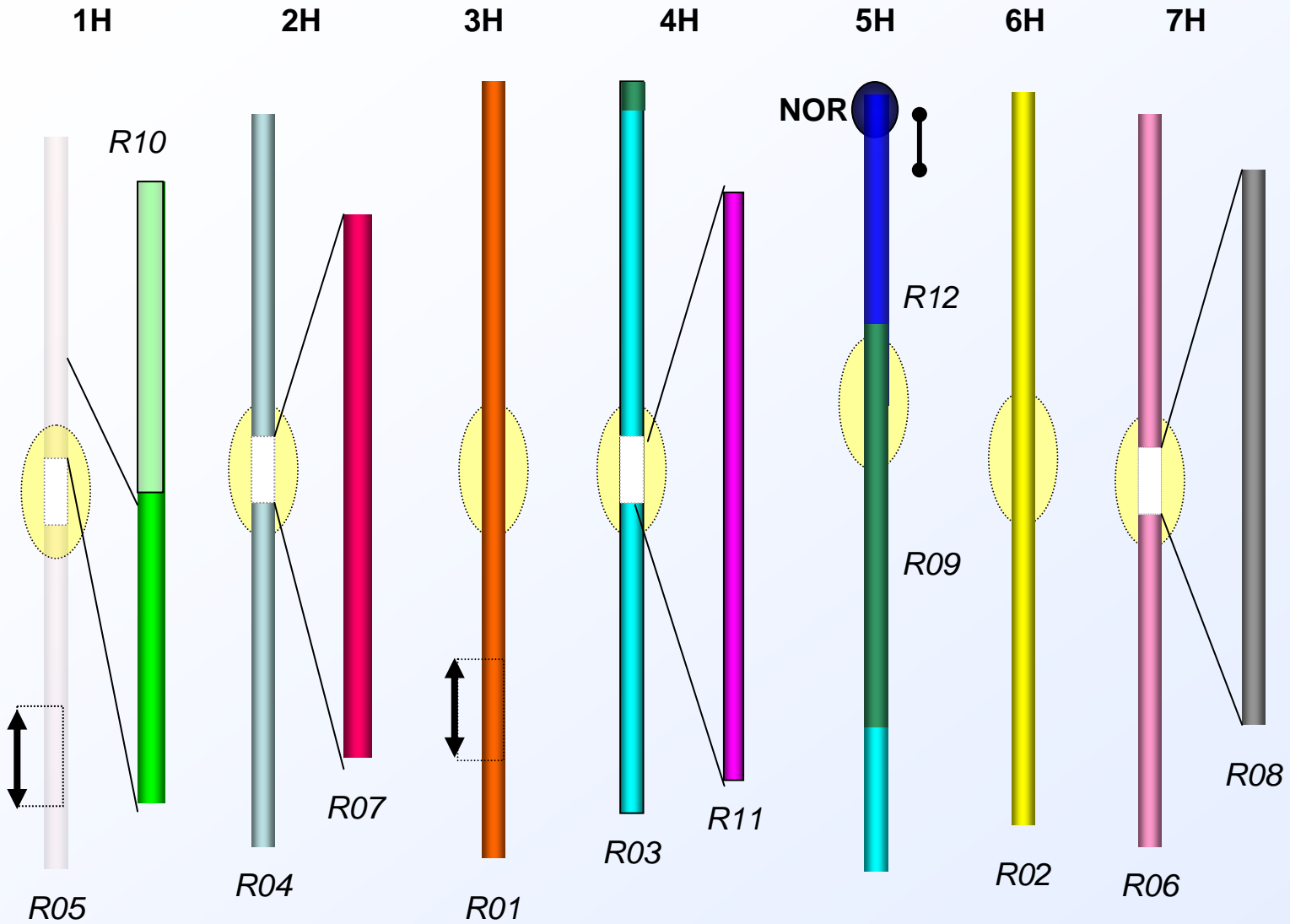


# Barley Illumina SNP Map – Feb06



# Coverage of Barley Map

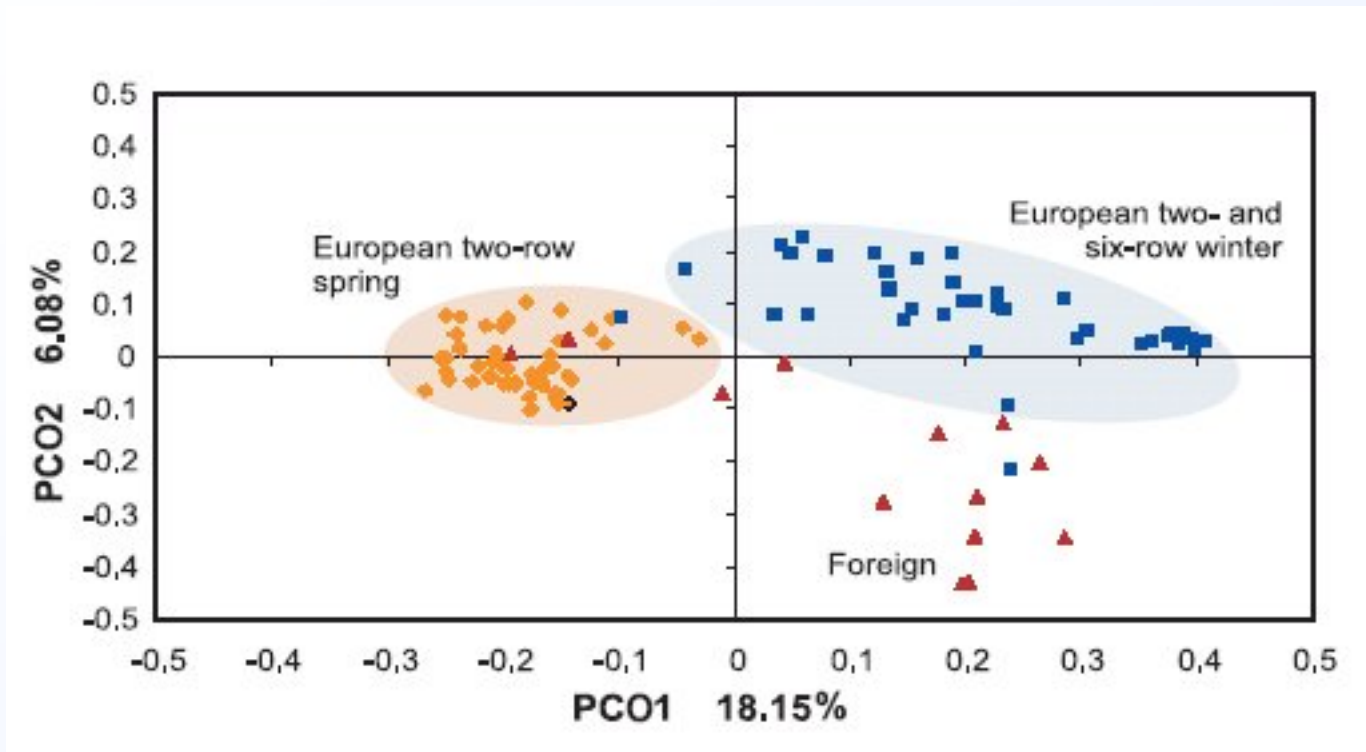




# Initial Diversity Studies

- Essentially single plate of barley accessions x Barley Illumina Pilot OPA1 (1536 barley SNPs)
- Total – 102 accessions
- Spring barley – 60 accessions (51 European)
- Winter barley – 42 accessions (38 European)
- European barley – 89 accessions
- Other material – parents of North American barley mapping populations (Steptoe, Morex, OWB D, OWB R, Dicktoo, Harrington), Japanese accessions (Haruna Nijo, Mokusekko 3), a Syrian landrace Cyrrhus and *Hordeum vulgare ssp. spontaneum* accession Mehola

# PCO plots of diversity patterns in barley



# Recombination across Barley 7H



# Acknowledgments

- **SCRI:**

Robbie Waugh

Bill Thomas

Sharon Mudie

Arnis Druka

Nicky Bonar

Katrin MacKenzie

Linda Milne

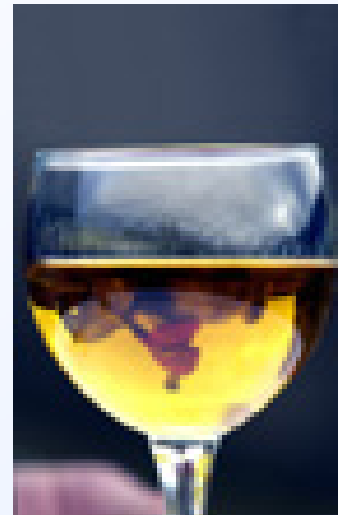
Paul Shaw

Luke Ramsay

Joanne Russell

Micha Bayer

Iain Milne



# Acknowledgements



Timothy Close

Jan Svennsson

Prasanna Bhat

Steve Wanamaker



UNIVERSITY OF  
BIRMINGHAM

Mike Kearsey

Zeweie Lou

Elena Potokina



Andreas Graner

Nils Stein

Rajeev Varshney

