Triticum dicoccoides в Хайфском Университете: Генофонд, структурная и функциональная геномика, клонирование генов, и приложения

http://evolution.haifa.ac.il

korol@research.haifa.ac.il/

ВИР, Санкт-Петербург, 5 октября 2011

The Institute of Evolution (IOE)

IOE was founded in 1976, by Prof. E. Nevo, now foreign member of the USA National Academy of Science.

IOE includes ~30 researchers; 11 are full or associated professors.

Research activities

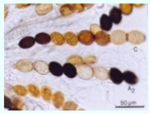
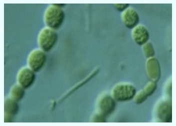
- Target organisms: microorganisms, plants, animals, humans
- Basic science is linked to applied research in agriculture, medicine, biotechnology, and industry
- Teaching at all levels – B.Sc., M.Sc., PhD, and posdocs
- Annual publication ~60-100 research articles, 2 books, 1-2 patents
- Grants: from BSF, BARD, GIF, DIP, CDR, INTAS, FP7, ISF

The Institute of Evolution

Research fields

- ☐ Biodiversity
- ☐ Molecular Biology & Evolution
- ☐ Evolutionary Cytogenetics
- ☐ Population Genetics
- ☐ Phylogenetics
- ☐ Genomics & Proteomics
- ☐ Bioinformatics
- ☐ Community Ecology
- ☐ Fungi Biotechnology
- ☐ Plant Ecology and Pollination
- ☐ Palaeobotany and Palaeoecology

Evolution & Biodiversity studies

- Ecology and genetics of natural populations
- Endangered species
- Evolution in action
- Assessment of environmental quality
- Genetic resources and their utilization






Wild cereals evolution & utilization program

- Origin of polyploid wheat genome
- Genetics of domestication
- Genetic resources: evaluation and conservation (Gene Bank)
- Cereal genomics: genetic/physical mapping, QTL analysis, MAS, cloning and sequencing

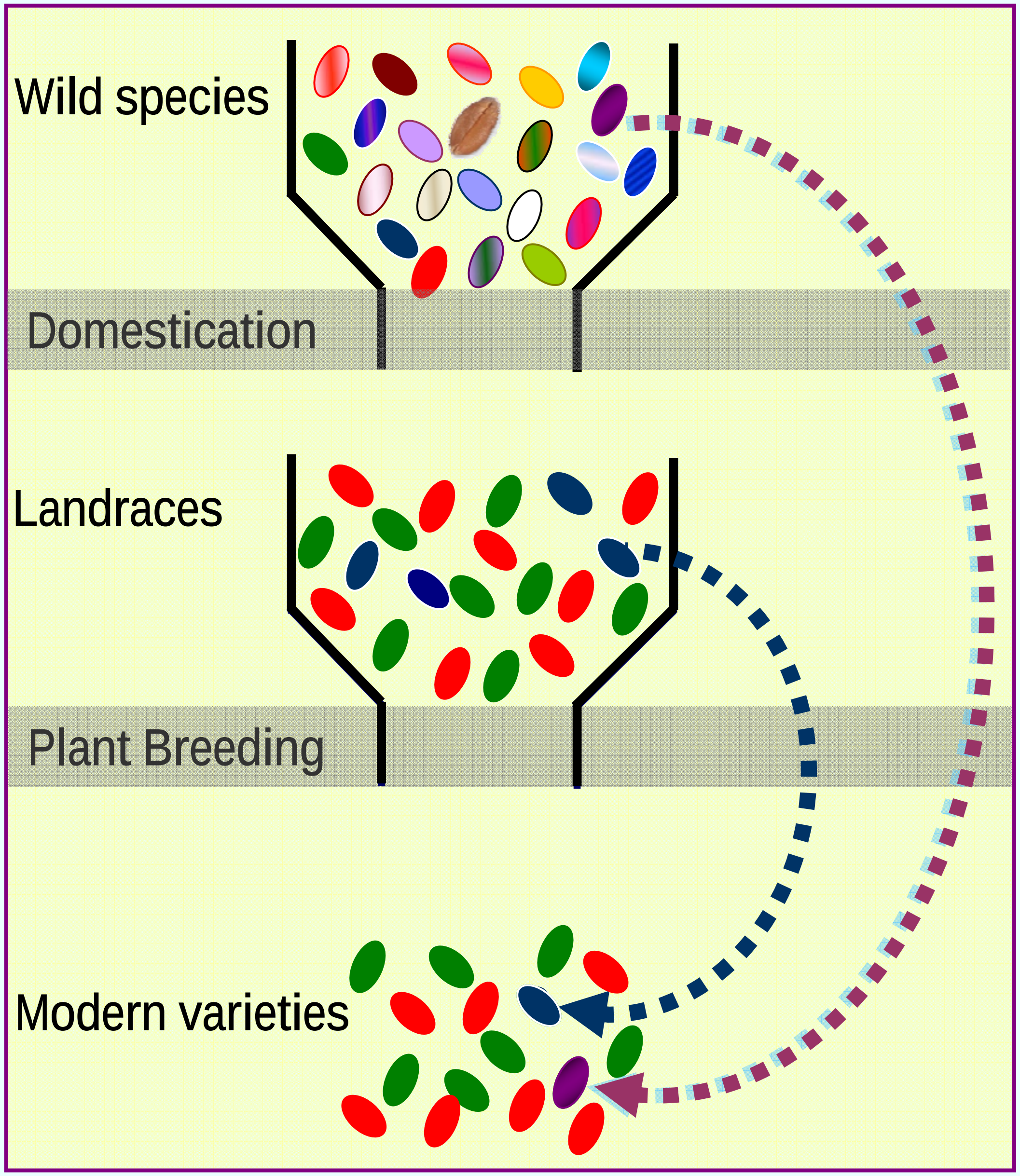
Wild progenitors of the Old World main domesticated plants

wild barley
H. spontaneum

wild emmer wheat
T. dicoccoides

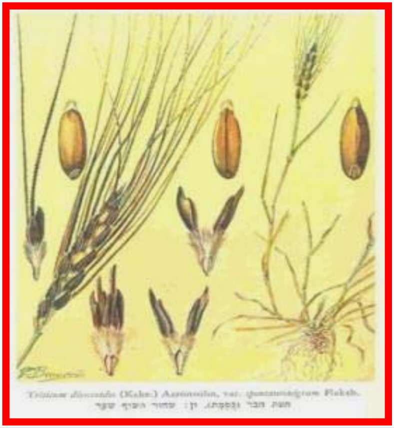
Erosion of Genetic Diversity in Crop Plants

- Domestication and plant breeding practices caused an erosion in the genetic diversity of crop plants.
- The most promising source for novel genes and alleles are the gene pools of wild relatives of cultivated crops.



Wild Emmer Wheat

Aaron Aaronsohn



- discovered in Israel in 1906
- is the progenitor of domesticated wheat

Triticum turgidum ssp. *dicoccoides* (Genome AABB)

Aaronsohn's vision was to use wild emmer wheat gene pool for improvement of cultivated wheat

100 year later ……

Wild Emmer Wheat

Can be crossed with

Spike shattering

T. durum

2n=4x=28, AABB, pasta

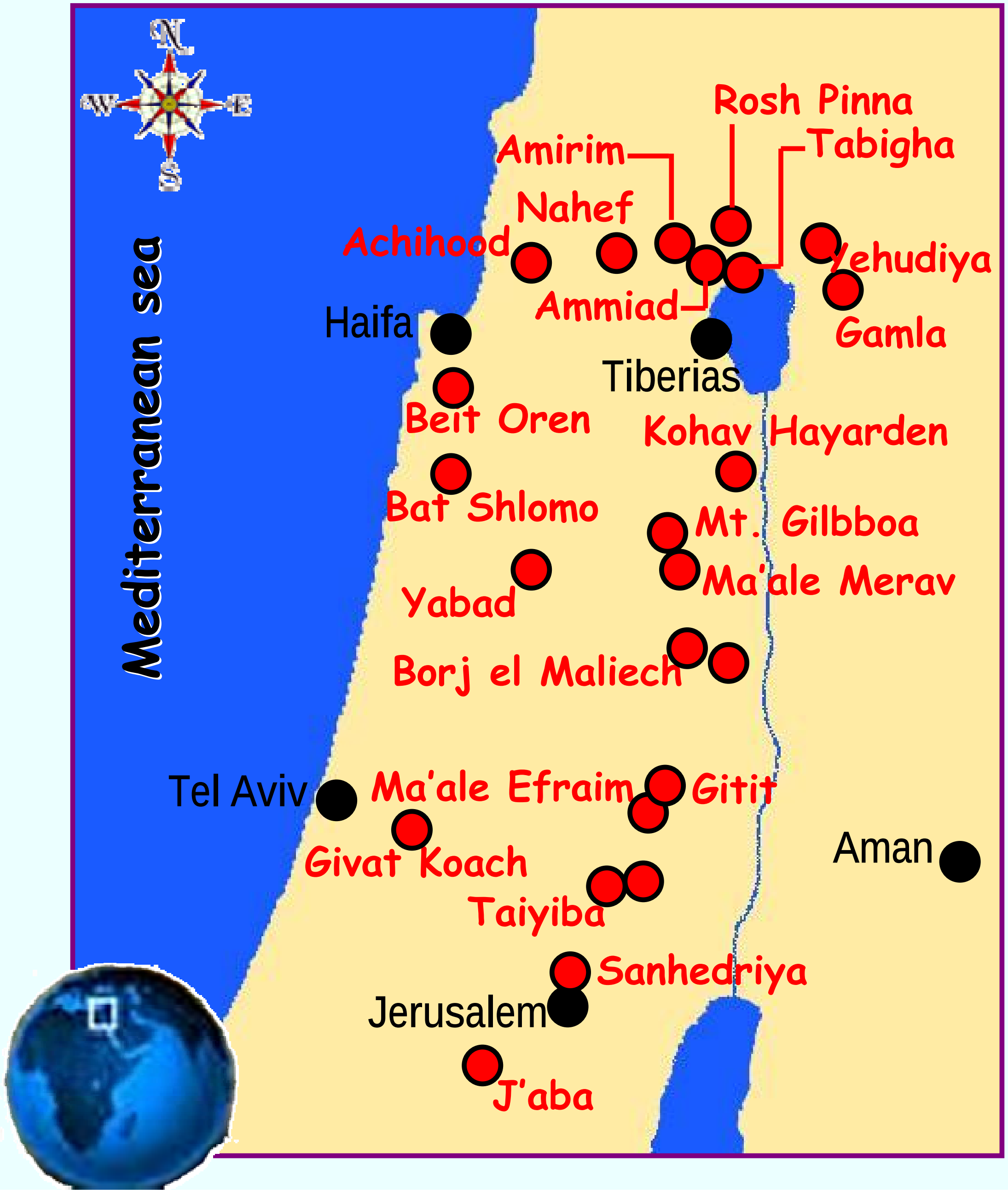
T. aestivum

2n=6x=42, AABBDD

Bread wheat

In the last 50 years about 50% of the wild wheat natural populations in Israel were lost due to urban/industrial development

Characterization of Wild Emmer Wheat Populations in Israel




Yehudiya, April, 2005

Agronomic traits

- Disease resistance
- Drought tolerance
- Grain proteins
- Grain minerals (Fe, Zn)
- Salt tolerance
- Herbicide resistance
- Photosynthetic yield

Biodiversity studies

Gene Bank of wild relatives of crop plants

- Wild emmer wheat, *Triticum dicoccoides* (from Israel, Turkey, Jordan)
- Wild barley, *Hordeum spontaneum* (from Israel, Turkey, and Iran)
- Wild wheat relative, *Aegilops* complex (170 populations of 11 species)
- Wild oat, *Avena sterilis* (1000 genotypes from Israel)
- *Brachypodium*, an important model cereal with a small genome
- Wild lettuce, *Lactuca serriola* (from Israel, USA, Europe; ~ 22 countries)
- Wild lettuce, *Lactuca saligna* (600 genotypes from Israel)
- Wild lettuce, *Lactuca aculeata* (180 genotypes from Israel, Turkey, Jordan)

The gene bank is used for **basic research** (genomics, population genetics and cytogenetics, domestication evolution) and enables coordinated and integrated sampling and **conservation**, and **breeding utilization** of wild germplasm.

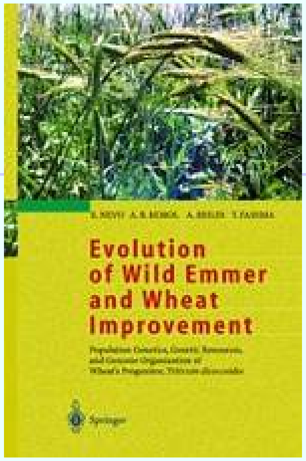
Wild Emmer Wheat Genetic Diversity

Agronomic traits

- Disease resistance
- Drought tolerance
- Grain proteins
- Salt tolerance
- Herbicide resistance
- Photosynthetic yield

Genetic Markers

- Allozymes
- SSRs
- RAPDs
- RGAs
- SNPs



Evolution
of Wild Emmer
and Wheat
Improvement

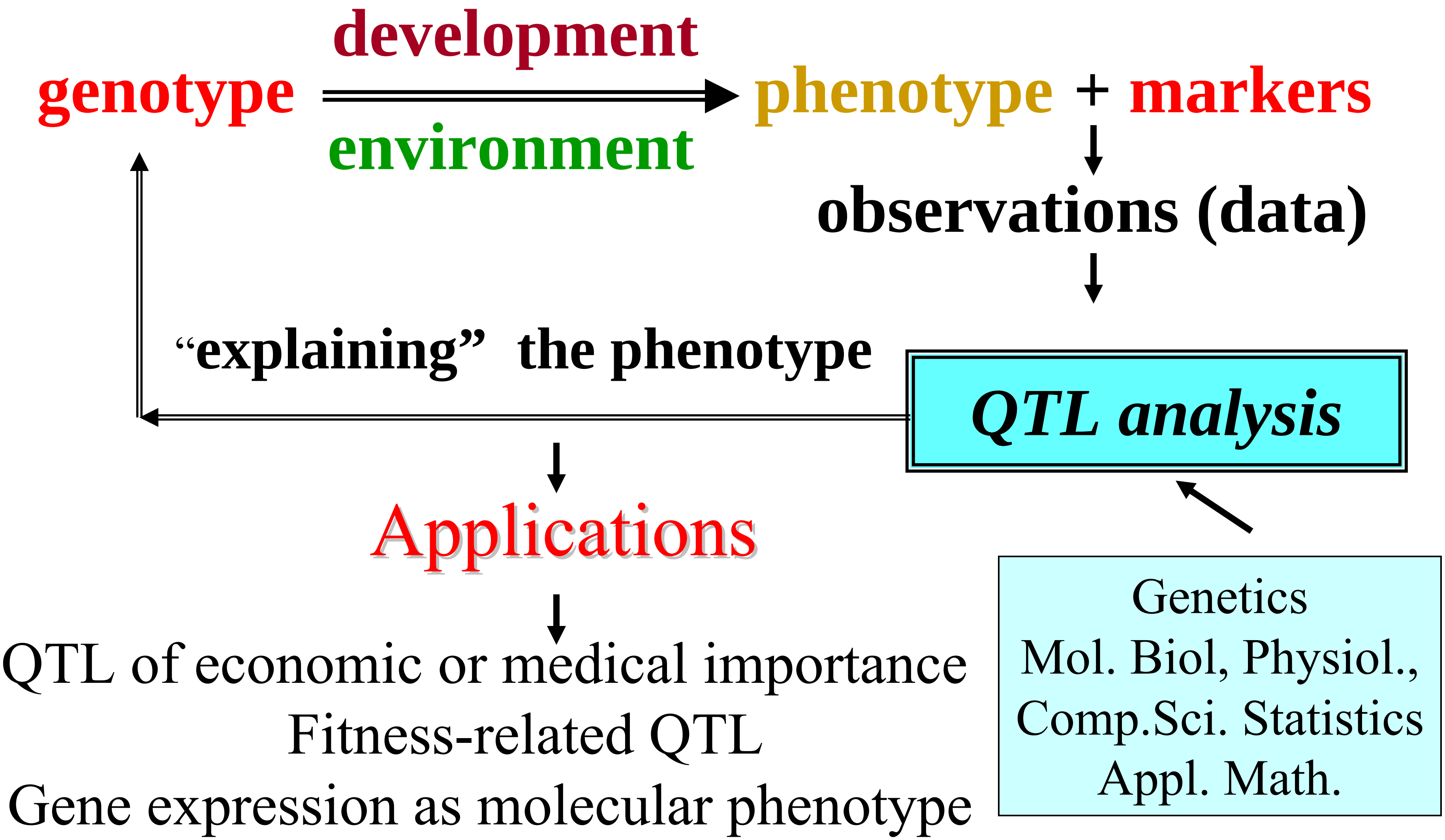
=► Rich diversity for crop improvement

Wheat Domestication Genomics

Study of genomic distribution of domestication QTLs
in a cross *T. durum* × *T. dicoccoides*

- How many loci (QTLs) involved
- How domestication QTLs are distributed
  - among genomes and chromosomes
  - within chromosomes
  - linkage vs. pleiotropy
- Characteristics of the QTLs

Genetic dissection of complex traits



genotype → (**development**, **environment**) → **phenotype** + **markers**

observations (data)

QTL analysis

“**explaining**” **the phenotype**

Applications

QTL of economic or medical importance
Fitness-related QTL
Gene expression as molecular phenotype

Genetics
Mol. Biol, Physiol.,
Comp.Sci. Statistics
Appl. Math.

About the map

Number of markers - 549

Total map length - 3170 cM

- A genome - 1589
- B genome - 1581

Average interval - 9.9 cM

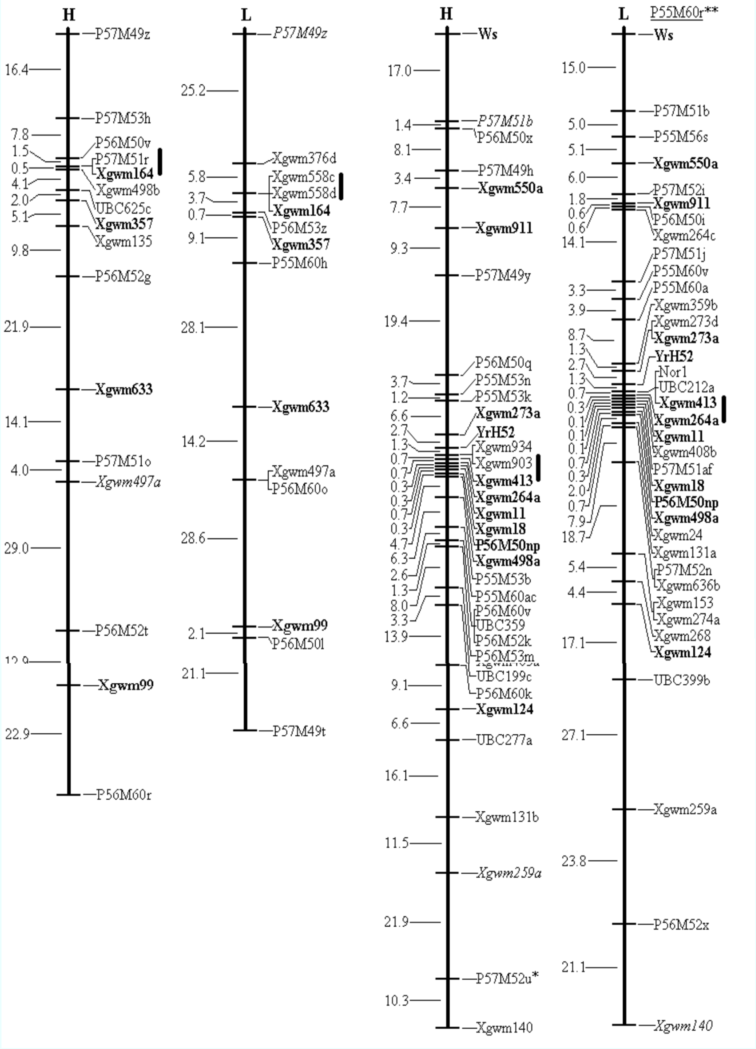
- A genome -11.1
- B genome - 8.9

Marker clustering (χ^2)

- A genome - 0.92 ns
- B genome - 19.97

Chr 1A

Chr 1B

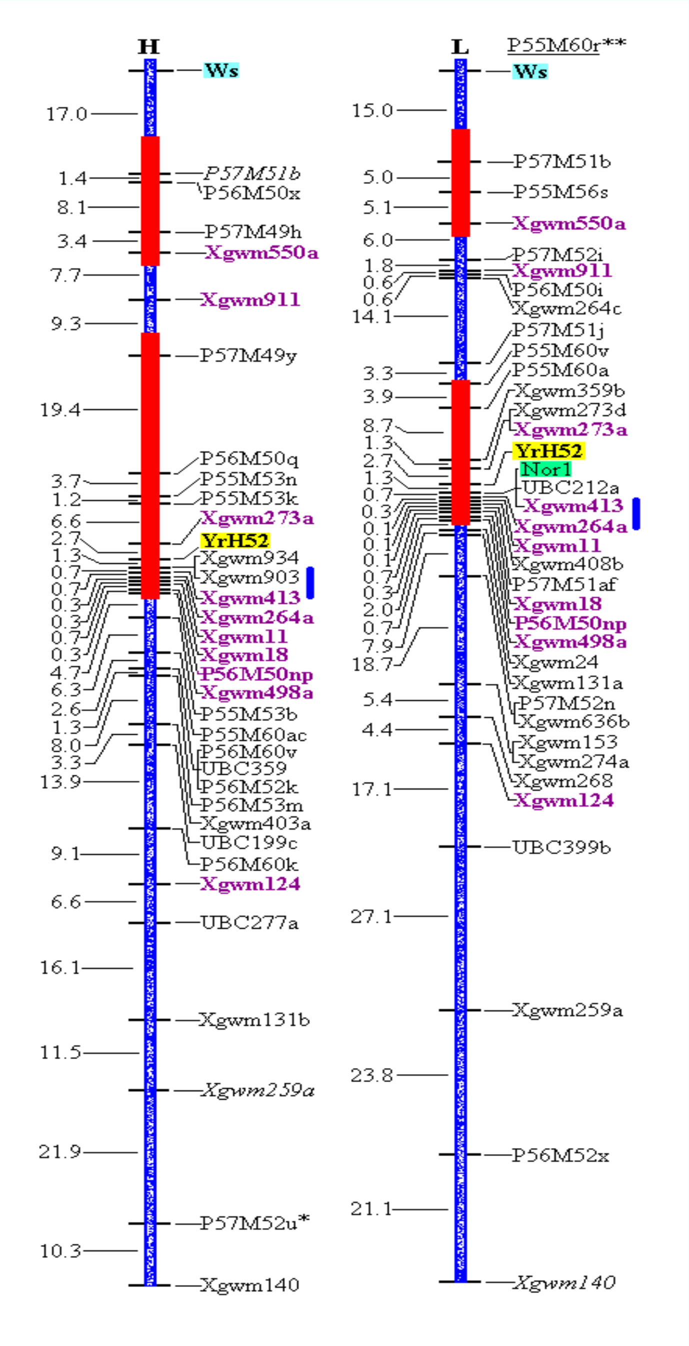


Peng *et al*. 2000, Genome Res. **10**:1509-31

Islands of *negative interference* in wheat chromosomes (e.g., 1B)

Negative interference seems to be not an exclusion (in wheat, barley, rye, *Drosophila*, and other species.

Together with marker reading errors, it may result in wrong marker orders → disagreement of genetic and physical maps



Highly pleiotropic *Domestication Syndrome* Factors (DSF) in tetraploid wheat

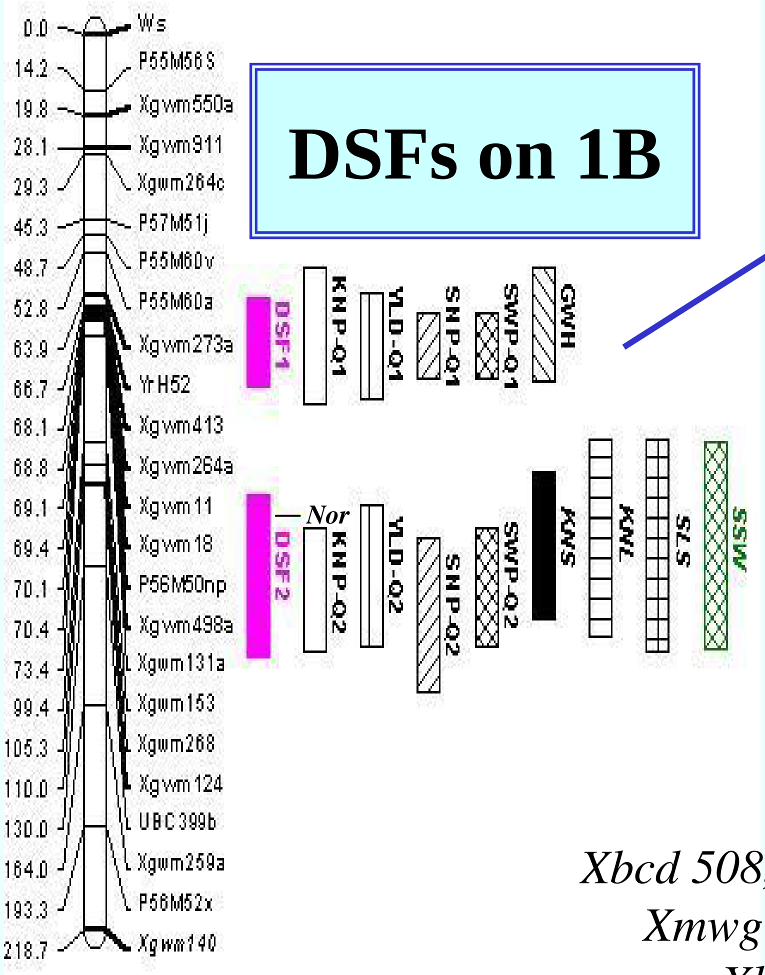
DSF	Chr.	Involved QTL effects	Interval	Position (cM)
DSF1	1B	KNP, YLD, GWH, SNP, SWP	Xgwm273a-Xgwm403a	68.6 ± 8.8
DSF2	1B	KNP, KNS, KNL, YLD, SNP, SLS, SWP, SSW	Xgwm124-P57M52u	133.9 ± 22.8
DSF3	2A	KNP, KNS,YLD, HD, SNP, SWP, SNP, SWP	Xgwm630c-Xgwm294	150.8 ± 31.4
DSF4	2A	KNP, KNS, KNL, YLD, GWH, SLS, SSW, SWP	Xgwm294-Br	216.8 ± 11.8
DSF5	3A	KNP, KNS, KNL, YLD, SSW, SWP	Xgwm218-Xgwm638	148.1 ± 28.4
DSF6	5A	KNP, KNS, KNL, YLD, HT, SLS, SSW, SWP	Xgwm154-P56M50m	72.5 ± 12.6
DSF7	5A	*all traits*	Xgwm186-P56M53c	144.2 ± 29.3

DSFs - *domestication syndrome* factors

Summary of the mapping results

- Seven intervals of strong effects (DSFs)
  - *domestication syndrome factors*
- Six out of seven DSFs appear as linked pairs
  - *some with same-sign* QTL *effects* (*e.g., on* **5*A***), *but most with opposite-sign effects* (*e.g., on* **2*A***)
- An excess of detected QTL effects in A genome
- Remarkably, DSFs coincide with gene-rich regions

Mapping wheat domestication syndrome factors



DSFs on 1B

Marker cluster

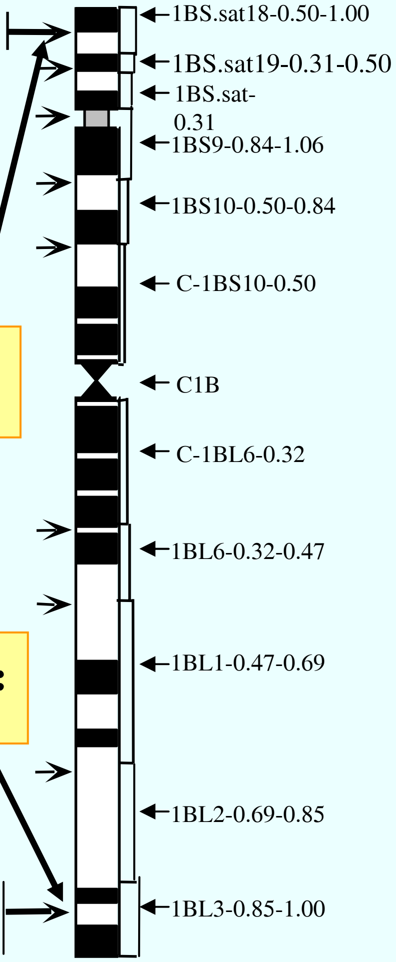
XksuE19, ***Nor***,
XksuD14a, ***XksuD14b***,
XksuD14c, *Xcdo388*,
Xmwg36, ***Xmwg938***,
Xbcd1434

168 ESTs*

68 ESTs*

Xbcd 508, *Adh*, *Xbcd304*, *XksuG34*,
Xmwg710, *Xbcd1562*, *Xwg241*,
XksuE11, *Xabg387*, *XksuI27*

Physical interval

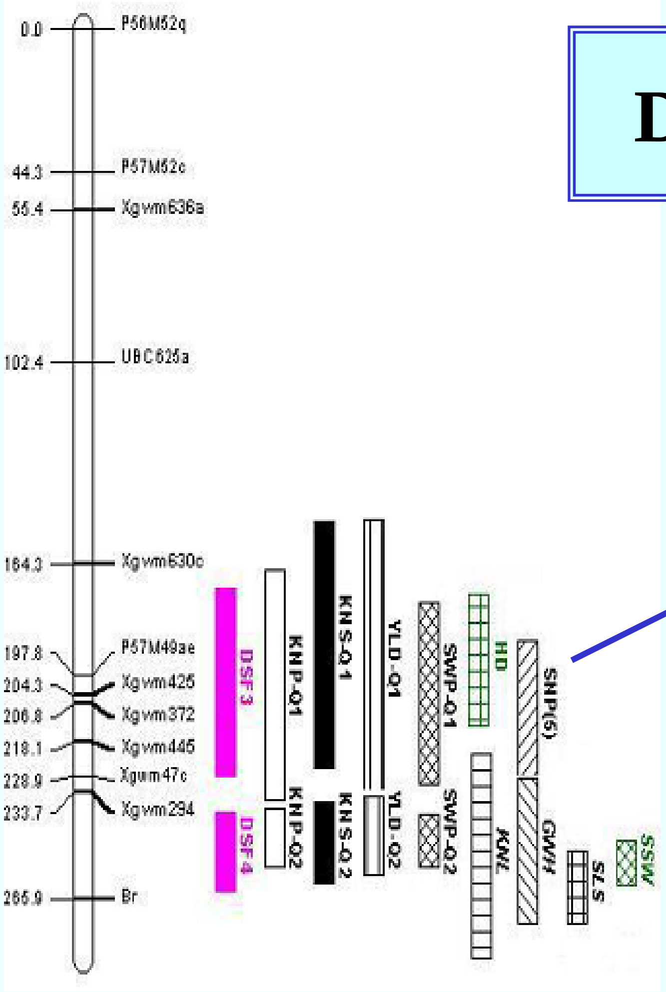


Peng *et al*. 2003
PNAS 100:
2489-2494

***EST loci mapped as of July, 2003**
http://wheat.pw.usda.gov/NSF/project/ mapping_data.html

Gill *et al*. 1996
Genetics 144:
1883-1891

DSFs on 2A



Physical intervals

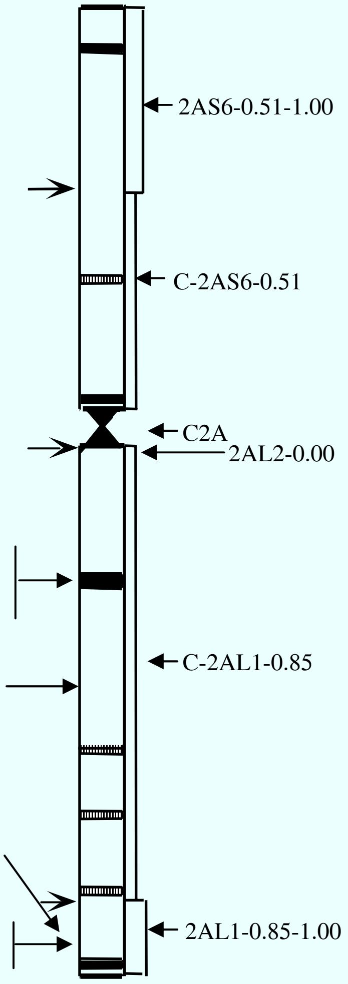
Marker cluster

Xpsr101, Xpsr102, Xpsr112,
Xpsr388, XksuF15,
XksuD22, XksuE16,
XksuF2, XksuD8, Xbcd135

305 ESTs*

104 ESTs*

XksuF41, ***XksuH16****,*
XksuI24, Xwg645

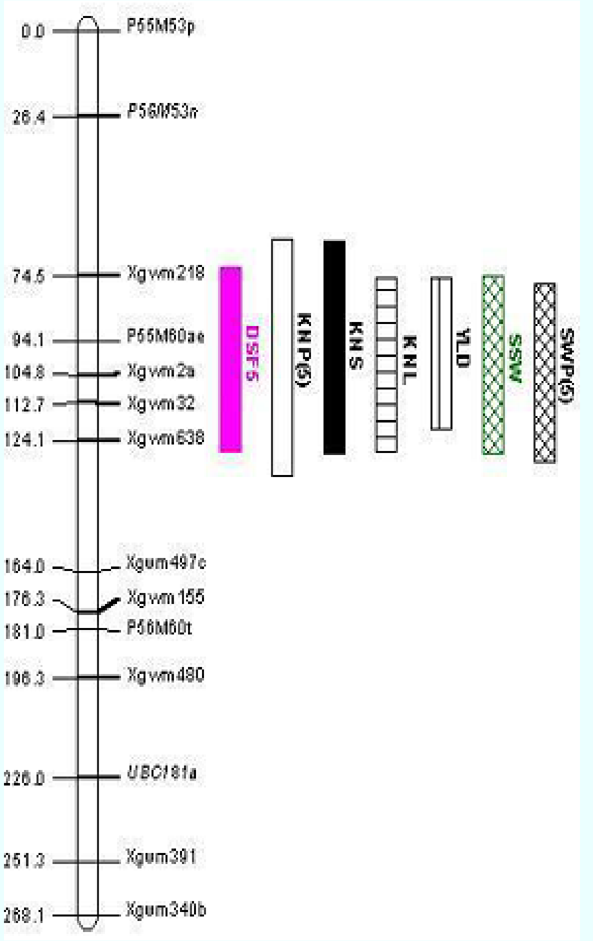


Peng *et al.*
2003 *PNAS*
100: 2489-94

*** EST loci mapped as of 2003**
http://wheat.pw.usda.gov/NSF/
project/mapping_data.html

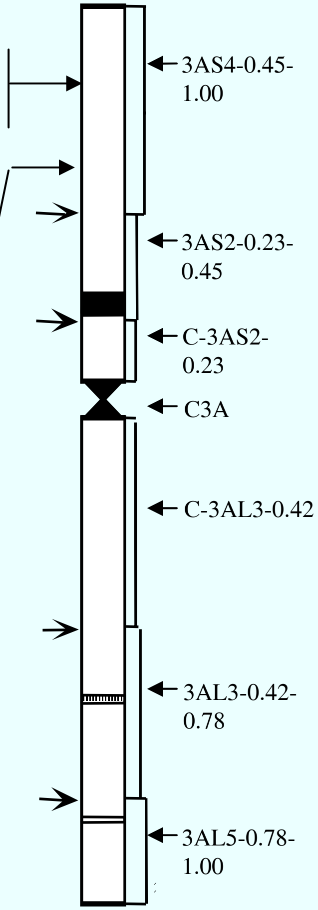
Delaney *et al.*
2003 *Theor Appl*
***Genet* 91: 568-73**

Marker cluster Physical intervals



XksuE2, XksuG13, XksuG53, XksuH7, XksuI32, Xpsr123, Xpsr598, Xpsr698, Xpsr903, Xpsr902, Xpsr910, Xpsr926, Xglk724

183 ESTs*



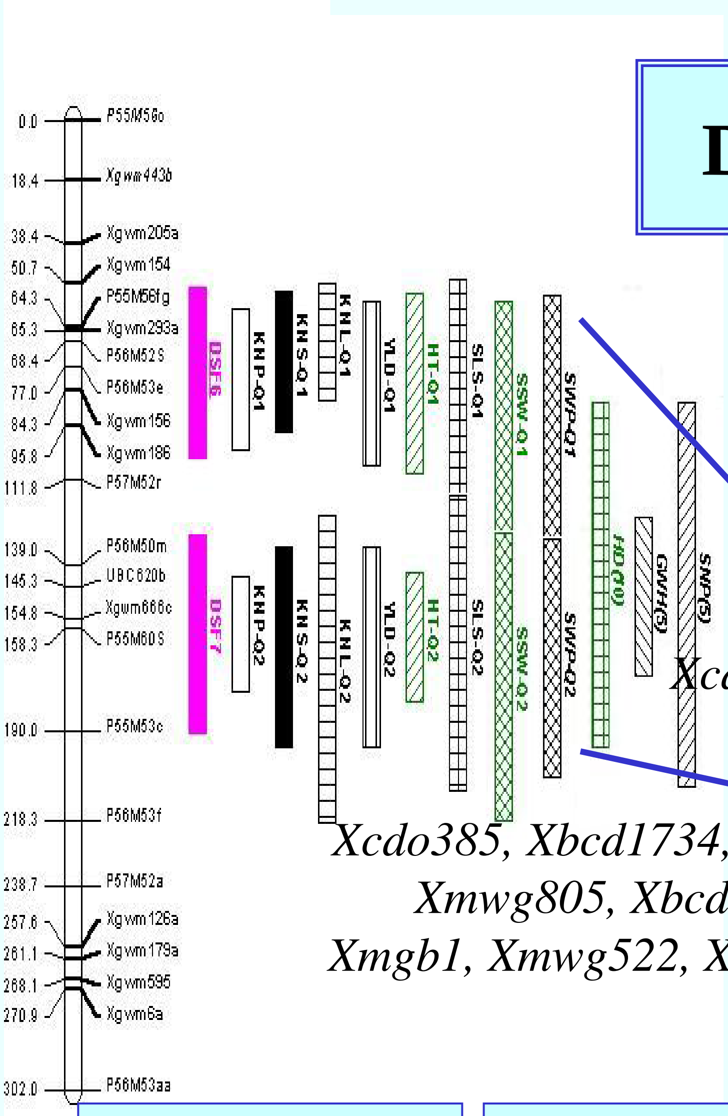
DSFs on 3A

Peng *et al.* 2003 PNAS 100: 2489-94

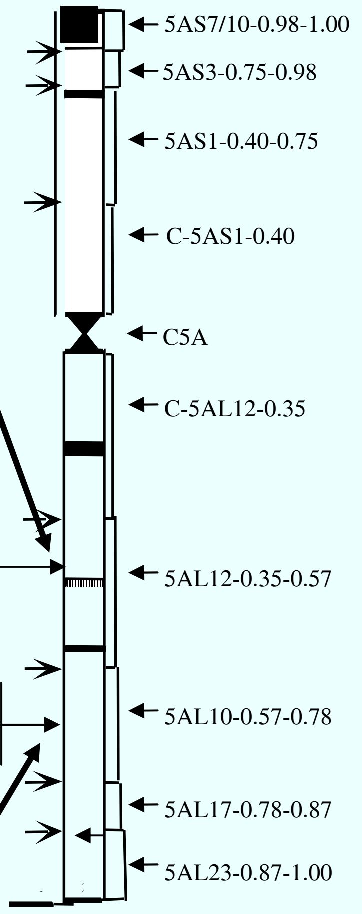
*** EST loci mapped as of 2003 http://wheat.pw.usda.gov/NSF/project/mapping_data.html**

Delaney *et al.* 2003 *Theor Appl Genet* 91: 780-2

DSFs on 5A



Physical intervals



122 ESTs*

Marker clusters

XksuG7, Xbcdb351, Xbcd157, Xcdo412, XksuH1, Xksu8,Xtag614, Xwg564, Xwg522, XksuQ45a

Xcdo385, Xbcd1734, Xbcd1235a, Xbcd21, Xmwg604, Xmwg805, Xbcd9a, Xcdo388, Xabc168, Xmgb63, Xmgb1, Xmwg522, Xbcd183, Xcdo1326, and other 60 markers

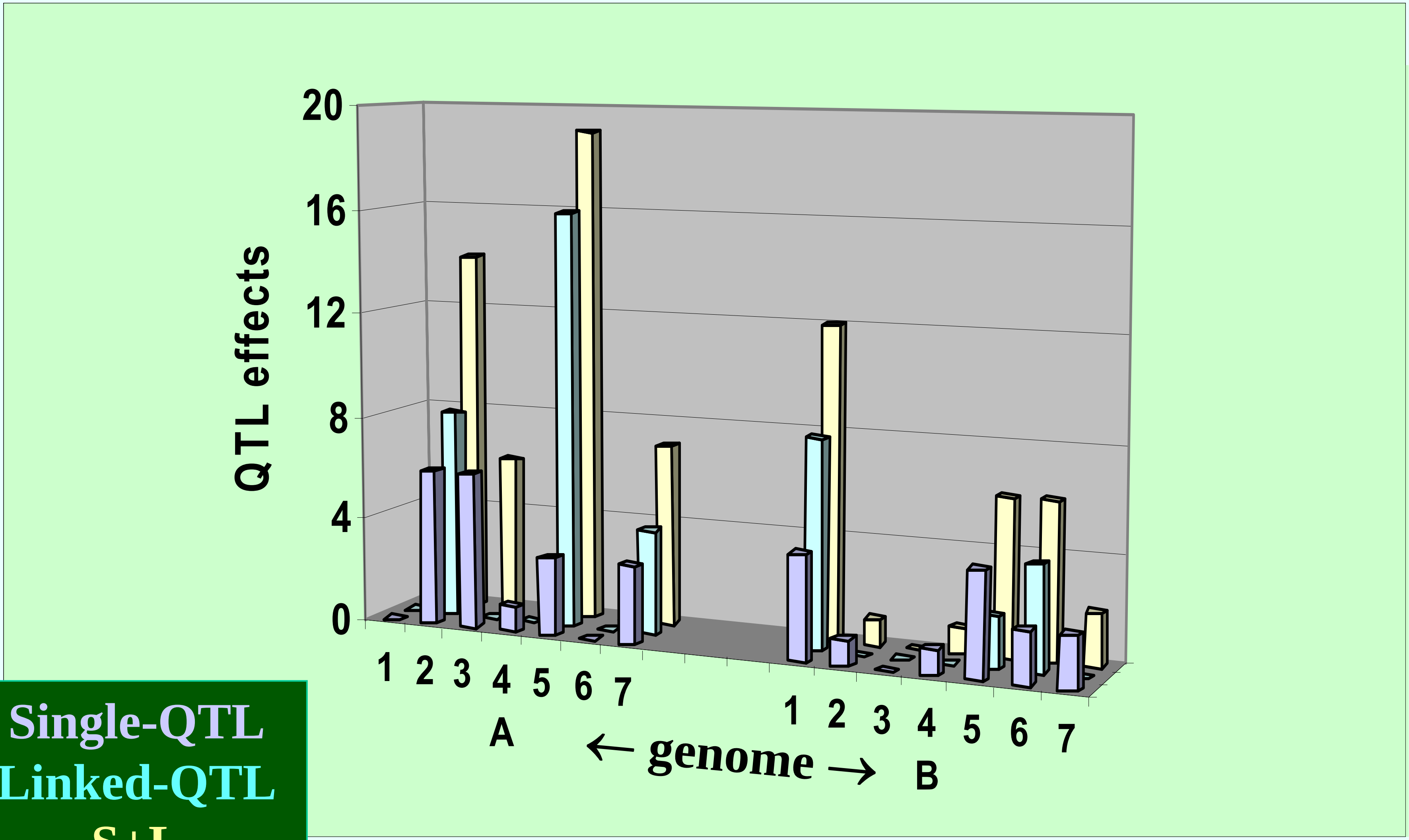
62 ESTs*

Peng *et al.*
2003 *PNAS*
100: 2489-94

*** EST loci mapped as of 2003**
http://wheat.pw.usda.gov/NSF/project/mapping_data.html

Faris *et al.* 2000
***Genetics* 154:**
823-835

Summary of the mapping results



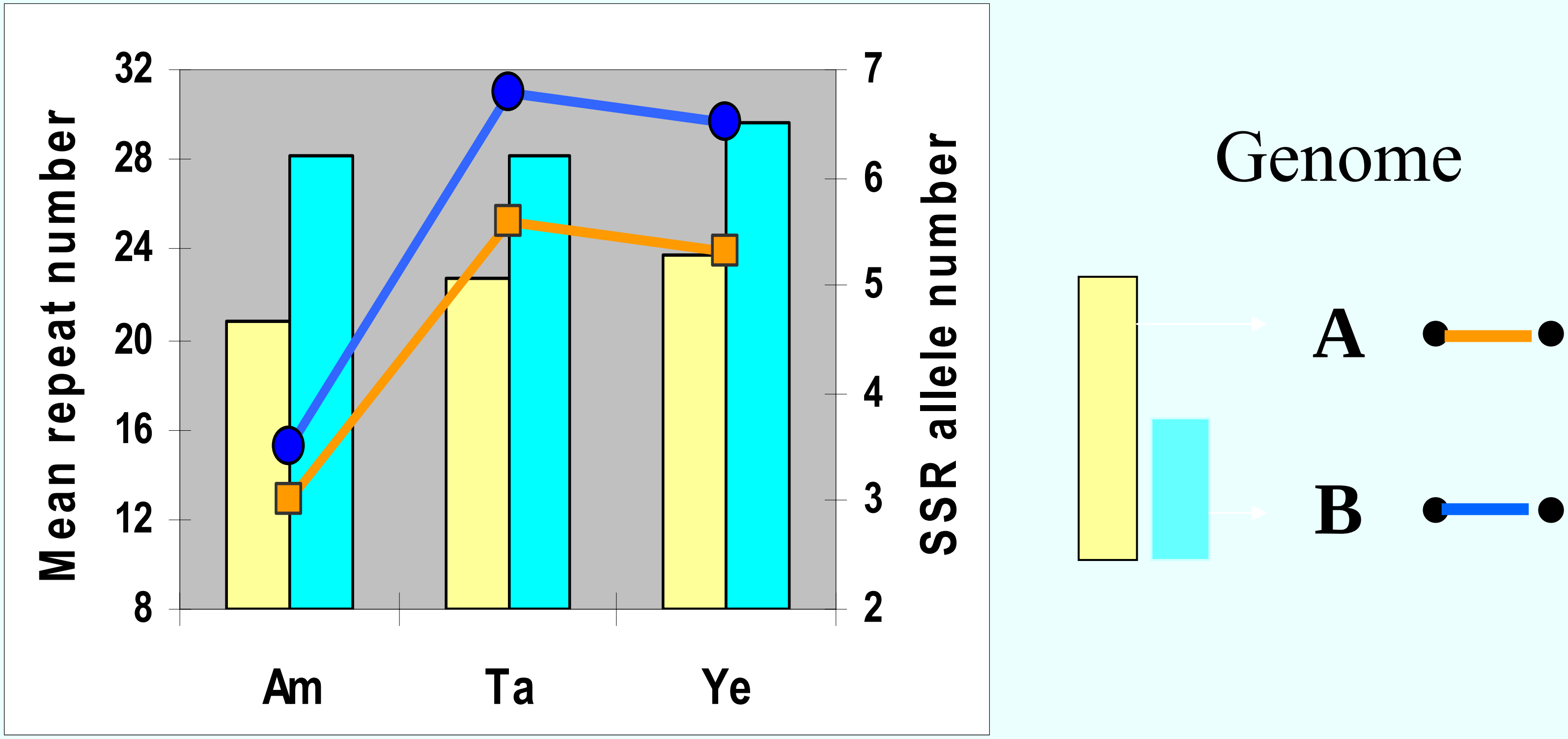
Single-QTL
Linked-QTL
S+L

Functional asymmetries in the tetraploid wheat genome

	A	**B**	***P*** $_{1:1}$
QTL effects (FDR 5%)	**37**	**16**	**0.004**
Disease resistance genes	**45**	**75**	**0.008**
Mapped ESTs ×10^3	**5.3**	**5.9**	**ns**
Mapped AFLP, %	**40**	**60**	**0.007**
Segregation-distorted loci	**23**	**8**	**0.0004**
Clustering of markers, χ^2	**0.92**	**19.9**	$<10^{-6}$

Genome asymmetry in population variation of wild wheat *T. dicoccoides*

(Li *et al*. 2000, *Mol. Biol. Evol.* 17: 851-862)



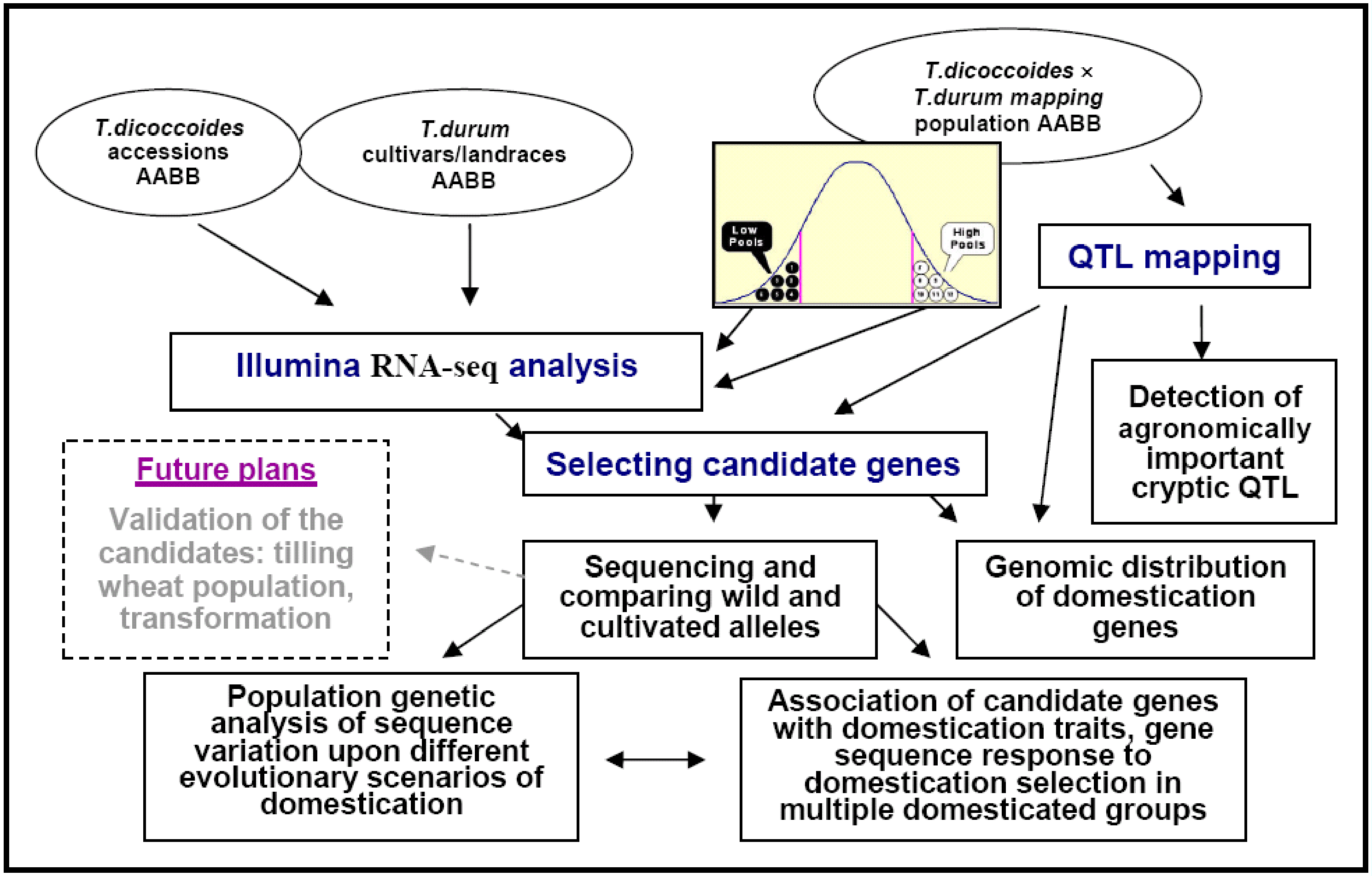
Am - Ammiad, Ta - Tabigha, Ye -Yehuddya

Hypothesis

In the wild, B genome plays a more important role in adaptation of the allopolyploid to ecological conditions than A. Thus, during domestication the response to artificial selection was provided by allelic variation at A genome loci.

It is a testable hypothesis, but needs a high resolution of QTL analysis complemented by functional genomics and sequencing candidate genes

Our new ISF project on wheat Domestication Genomics



T.dicoccoides accessions AABB

T.durum cultivars/landraces AABB

T.dicoccoides × *T.durum mapping* population AABB

QTL mapping

Illumina RNA-seq analysis

Detection of agronomically important cryptic QTL

Selecting candidate genes

Future plans

Validation of the candidates: tilling wheat population, transformation

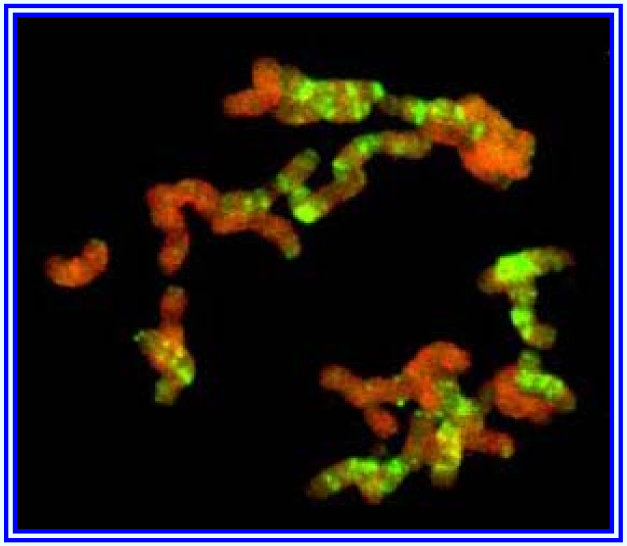
Sequencing and comparing wild and cultivated alleles

Genomic distribution of domestication genes

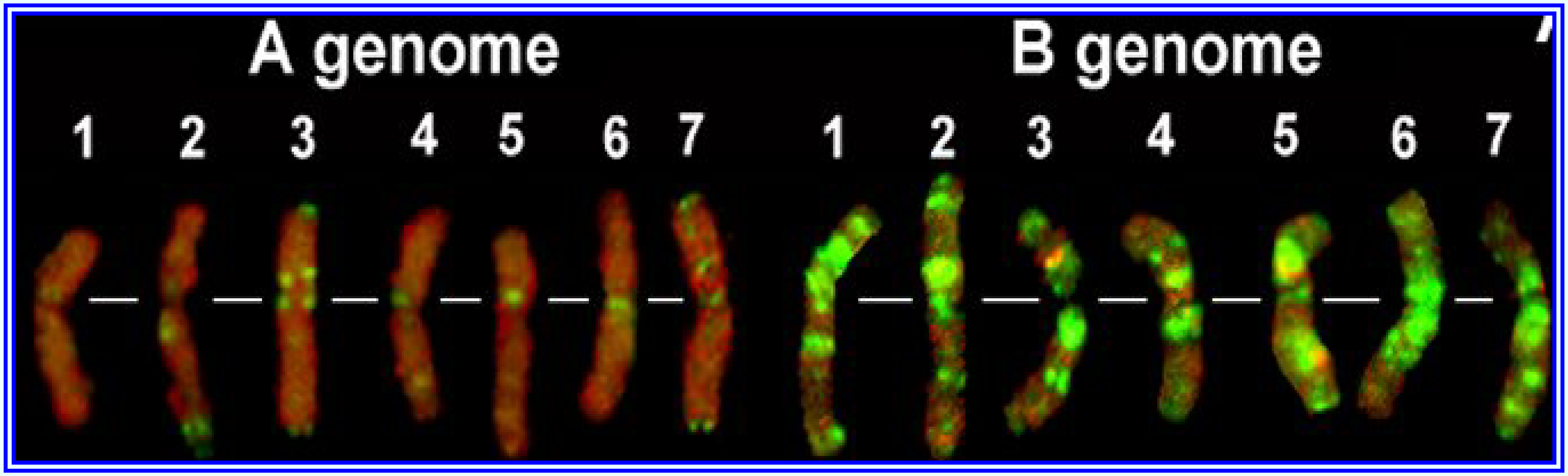
Population genetic analysis of sequence variation upon different evolutionary scenarios of domestication

Association of candidate genes with domestication traits, gene sequence response to domestication selection in multiple domesticated groups

Evolutionary Cytogenetics

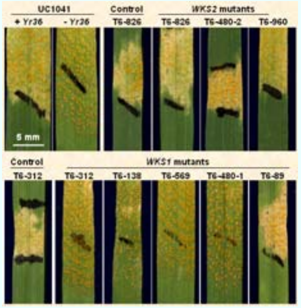


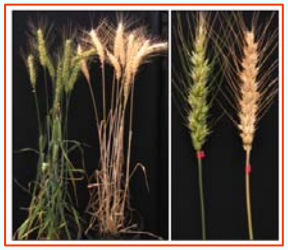
Genome in situ hybridization enables revealing and scoring evolutionary divergence of genomes as well as <u>introgressions</u> caused by interspecific recombination



Some Applied Wheat Genomic Studies

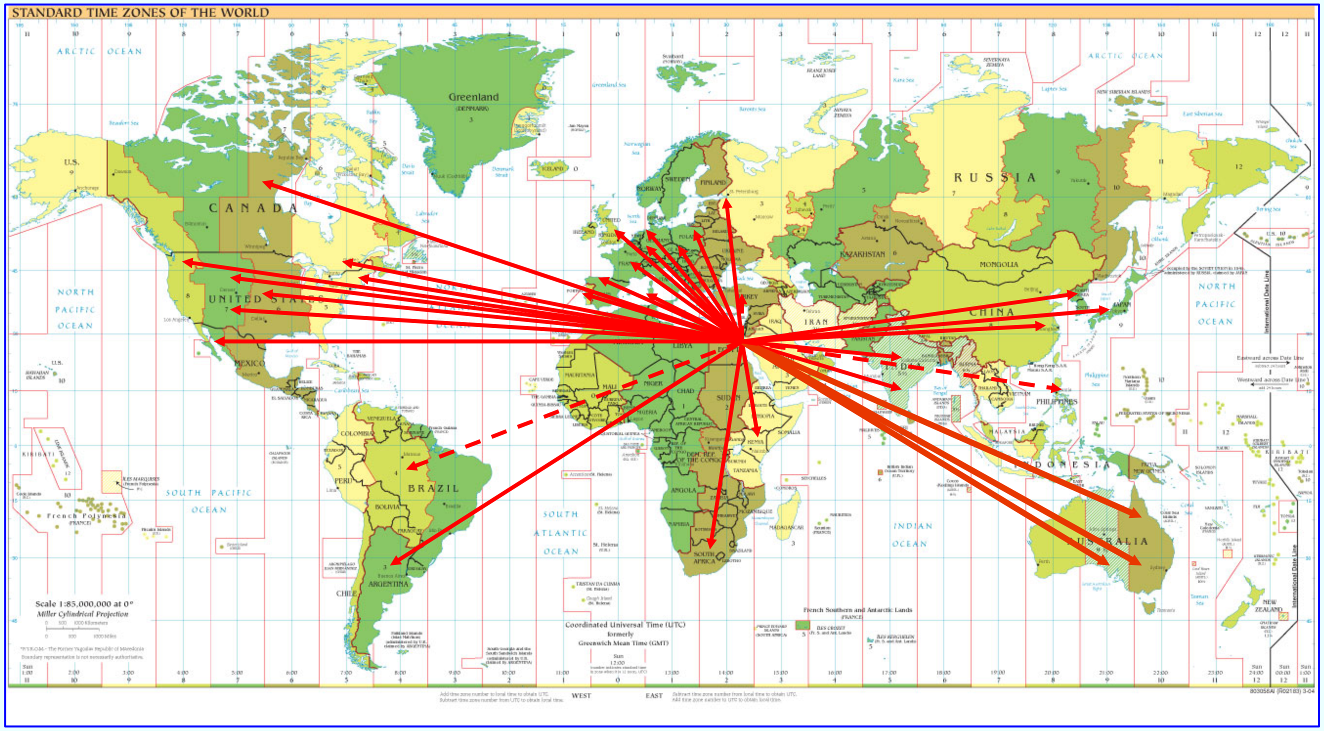
- **Software: Multi-Point, Multi-QTL and LTC - powerful tools for Genetic and Physical Mapping that enabled us to develop better maps of the complex wheat genome**
- ***Yr15, YrH52, PmG3M, Gpc-B1*** **and** ***Yr36*** **- novel genes derived from wild emmer wheat**
- ***Gpc-B1*** **was the first QTL to be cloned in wheat**
- ***Yr36*** **(and** ***Lr34*****) were the first quantitative resistance genes cloned in wheat**
- **The cloning of these genes is emphasizing the importance of wild wheat gene pool for crop improvement**





Our mapping tools are distributed worldwide

USA, Canada, UK, France, Germany, Italy, Spain, Poland, Holland, Portugal, Czech Republic, Belgium, Switzerland, Russia, India, Japan, China, South Korea, Kenya, South Africa, Argentina, Australia, Israel

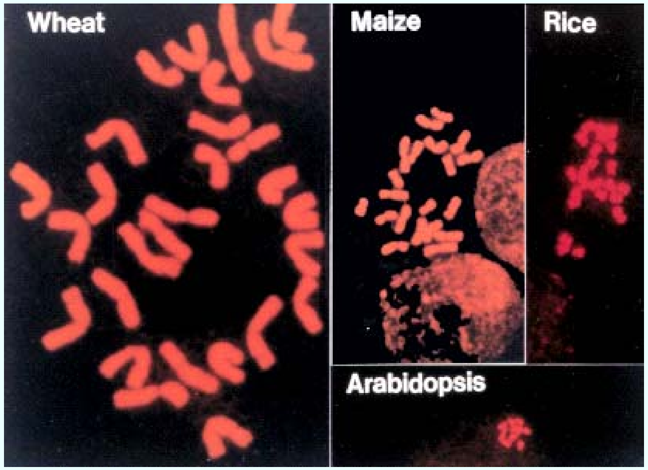


Positional cloning of agriculturally important genes derived from wild emmer wheat

Gene	Trait	Type	Location	Current status
Gpc-B1	High grain protein and mineral content	QTL	6BS	Cloned **Uauy et al. *Science* 2006**
Yr36	High-temperature stripe rust resistance	QTL	6BS	Cloned **Fu et al. *Science* 2009**
Yr15	Stripe rust resistance	Major gene	1BS	Physical mapping
YrH52	Stripe rust resistance	Major gene	1BS	Fine mapping/ physical mapping
PmG3M	Powdery mildew resistance	Major gene	6BL	Fine mapping

Complexity of Positional Cloning in Wheat

Relative Genome Size



(Bikram Gill, personal comm.)

1. Genome size 17,000 Mbp

Arabidopsis 145 Mbp (110×)
Rice 430 Mbp (35×)
Tomato 900 Mbp (16×)
Human 3,500 Mbp (5×)

2. Polyploidy

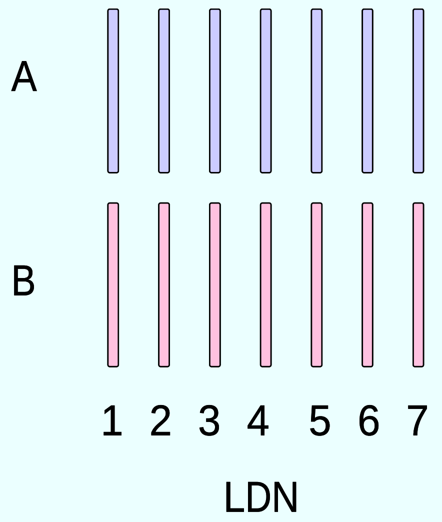
Triticum aestivum (6x) AABBDD
Triticum dicoccoides (4x) AABB

95% similarity between A, B & D genomes

3. Repetitive sequences

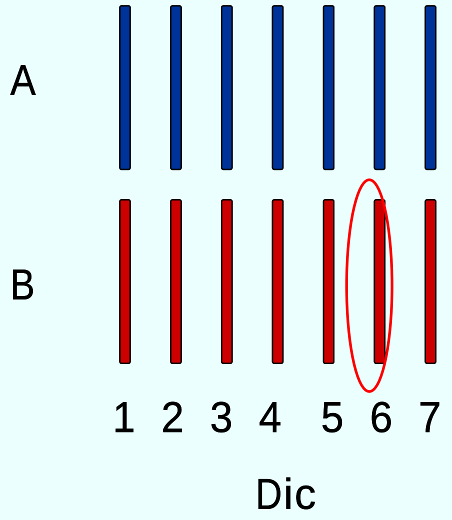
More than 80% repetitive (mainly transposon elements)
Wheat Genome is not sequenced (yet !!!)

A QTL for **Grain Protein Content** was identified using a set of Chromosome Substitution Lines



T. turgidum
ssp. *durum*

Low GPC



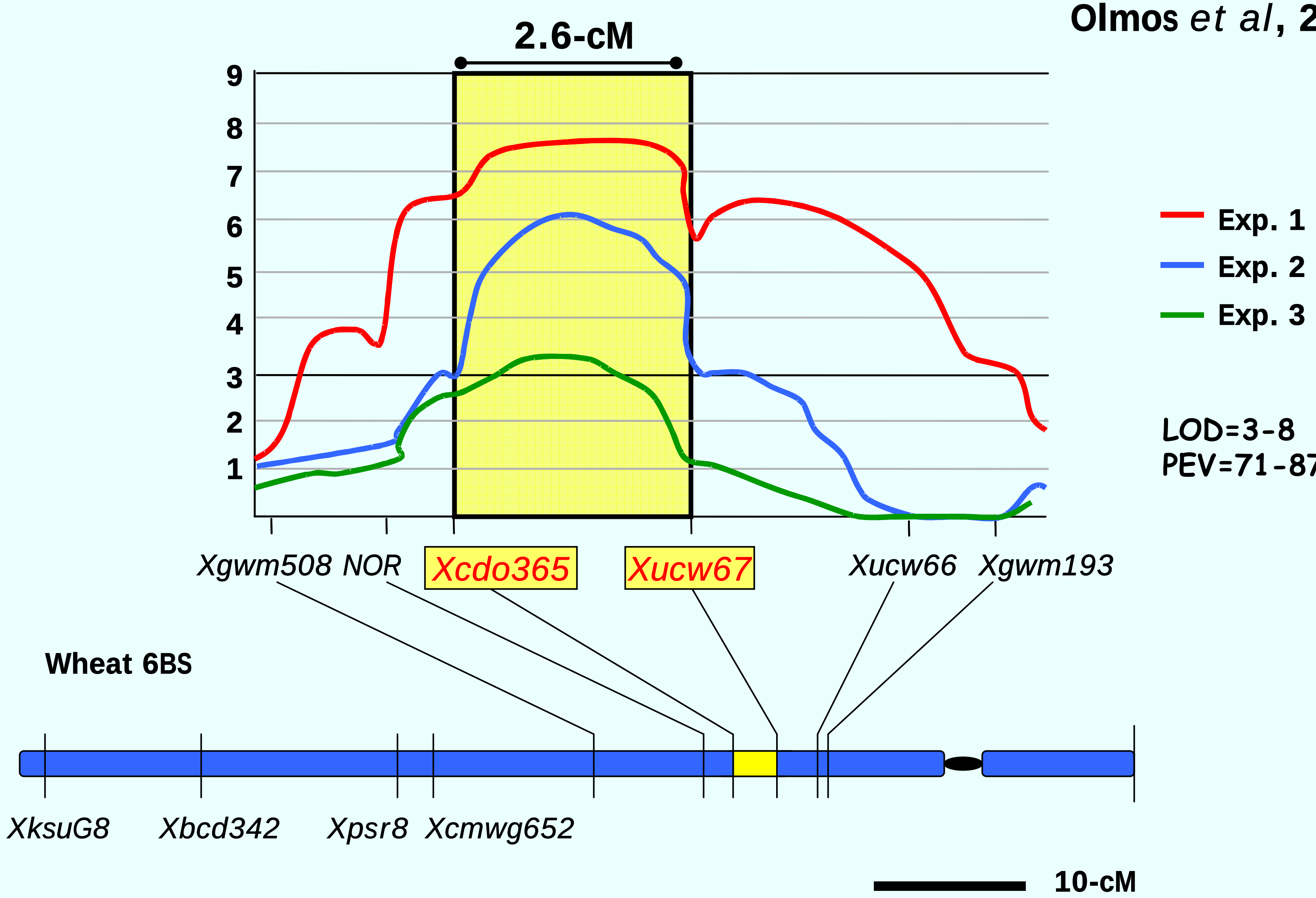
T. turgidum
ssp. *dicoccoides*

High GPC

(Joppa *et al.*, 1990)

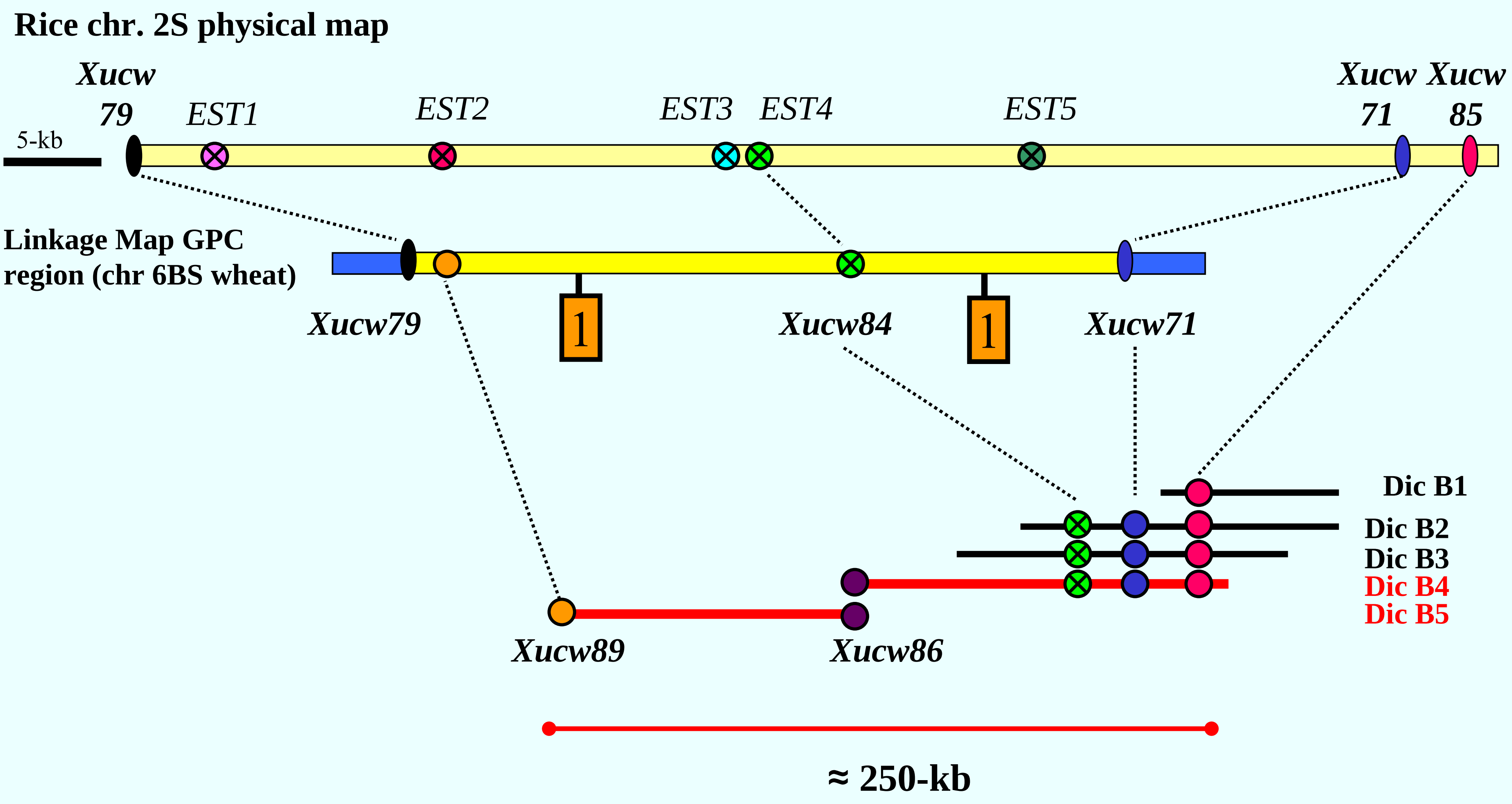
QTL mapping of *Gpc-B1* on 6BS

Olmos *et al*, 2003



LOD=3-8
PEV=71-87%

Complete Physical Map of *Gpc-B1*



Distelfeld et al. (2006) New Phytologist

Proposed model for *Gpc-B1* effects

Earlier Senescence

Degradation of cell components

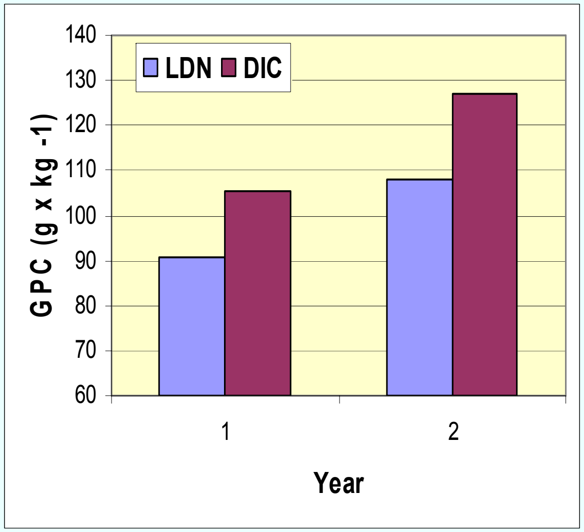
Higher levels of soluble proteins and minerals available for remobilization

More efficient remobilization

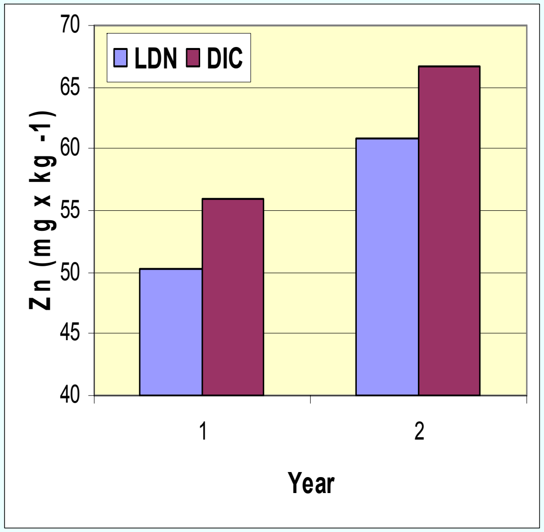
Higher grain protein and mineral concentrations (N, Zn, Fe)

LDN DIC

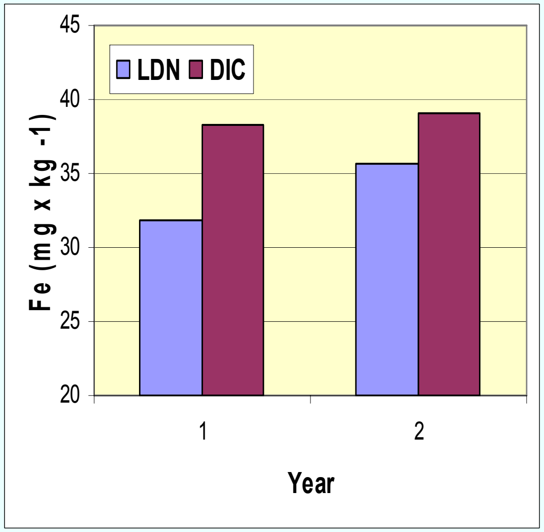
QTL effects of *GPC-B1* on:
Protein, Zn and Fe concentrations in the grain



GPC



Zn



Fe

Average increase over 2 years of:

18 % in GPC

11 % in Zinc

15 % in Iron

Distelfeld et al. 2007, In collaboration with I.Cakmak & H.Budak

Importance of *GPC-B1* for wheat nutritional value

➢600 million tons of wheat per year

⬇

Produce more than 60 million tons of protein/year

Zinc deficiency in diet can affect a range of functions:

- Immune system
- Growth rate
- Brain development
- Sexual Reproduction

QTL Mapping of Valuable Traits in Tetraploid Wheat

Gitit × LDN Mapping Population
(152 F_6 Recombinant Inbred Lines)

Gitit 18-16
Wild emmer wheat

×

Langdon
Durum wheat

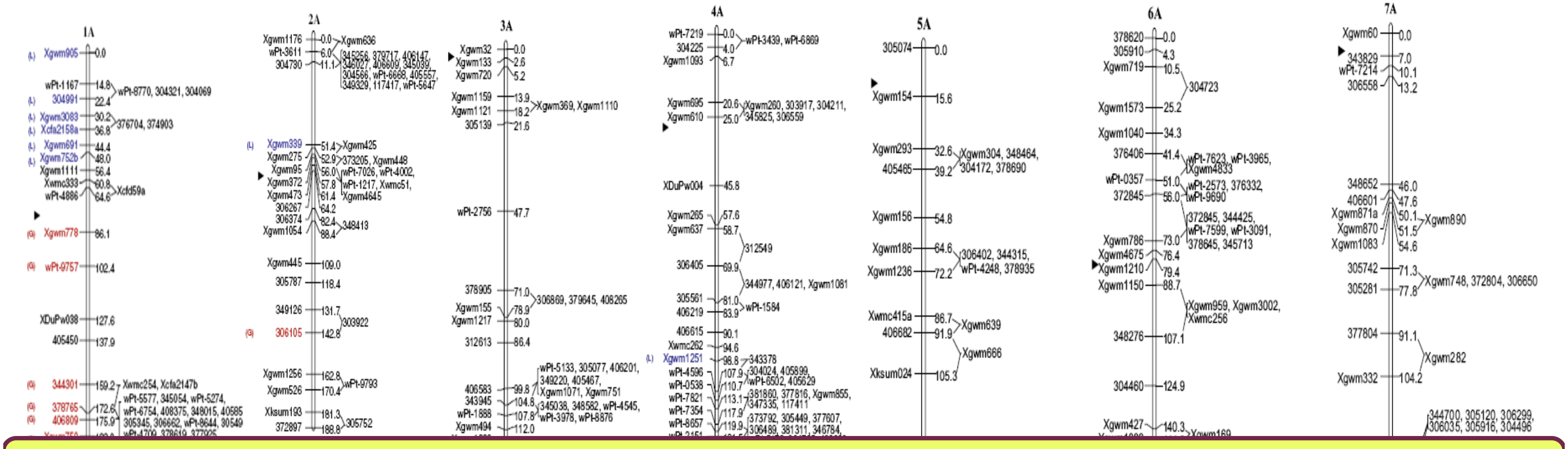
152 F_6 RILs

Spike Morphology

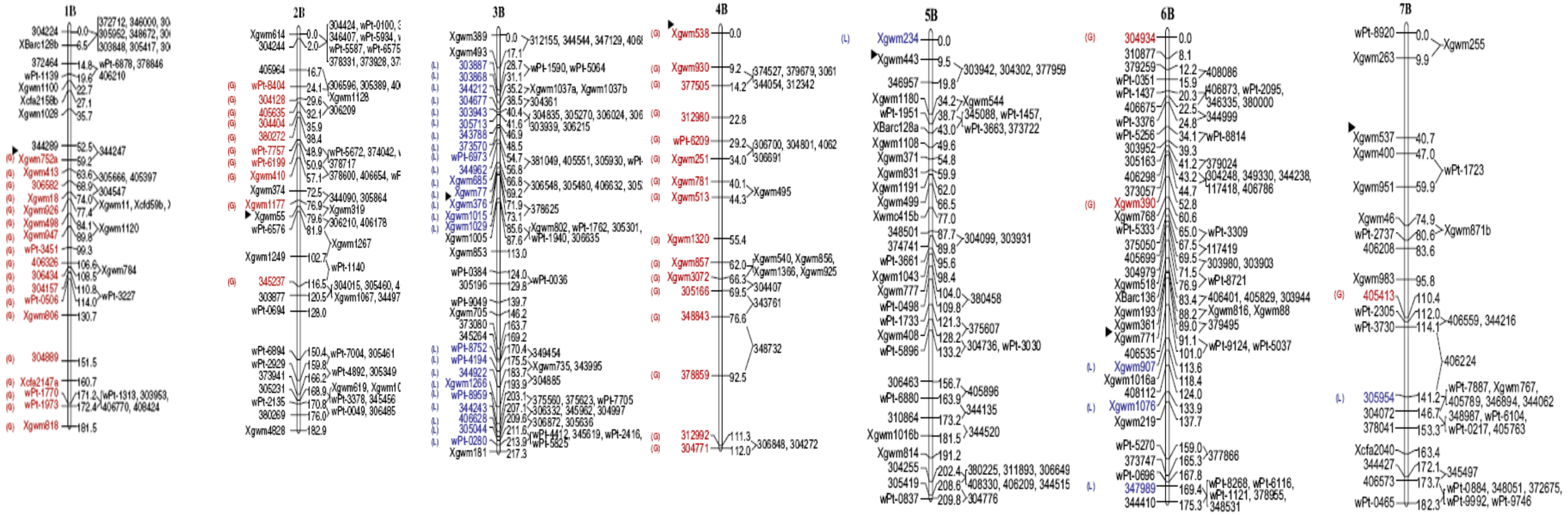
- Color
- Length
- Seeds per spike
- Shape
- Weight
- etc…….

Mineral Content
Drought tolerance

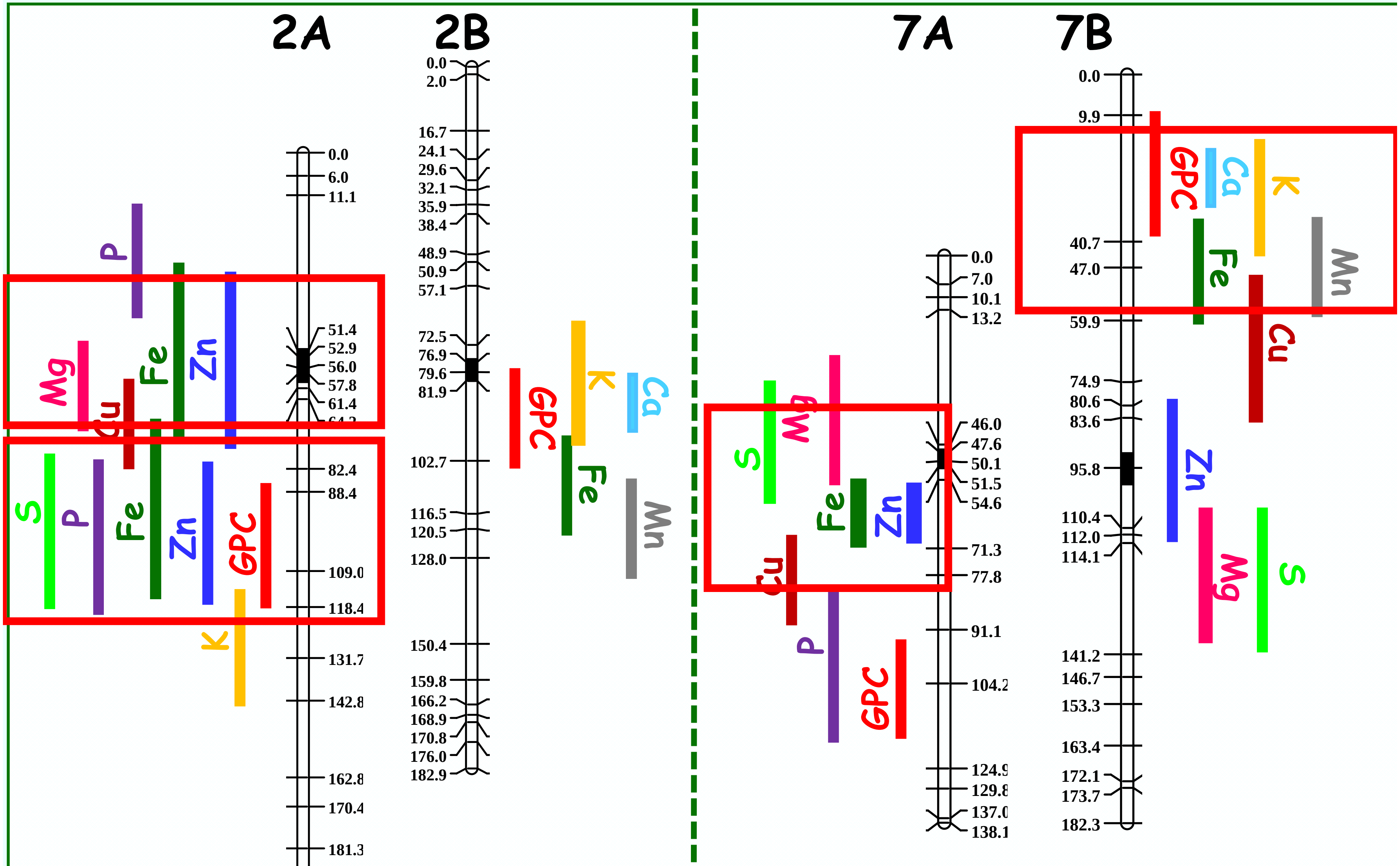
Genetic map based on 600 DNA markers



2300cM long, 7.5cM/Marker



QTL map of grain nutrients



Peleg *et al.* TAG (2009)

Emmer wheat as a source for drought stress tolerance

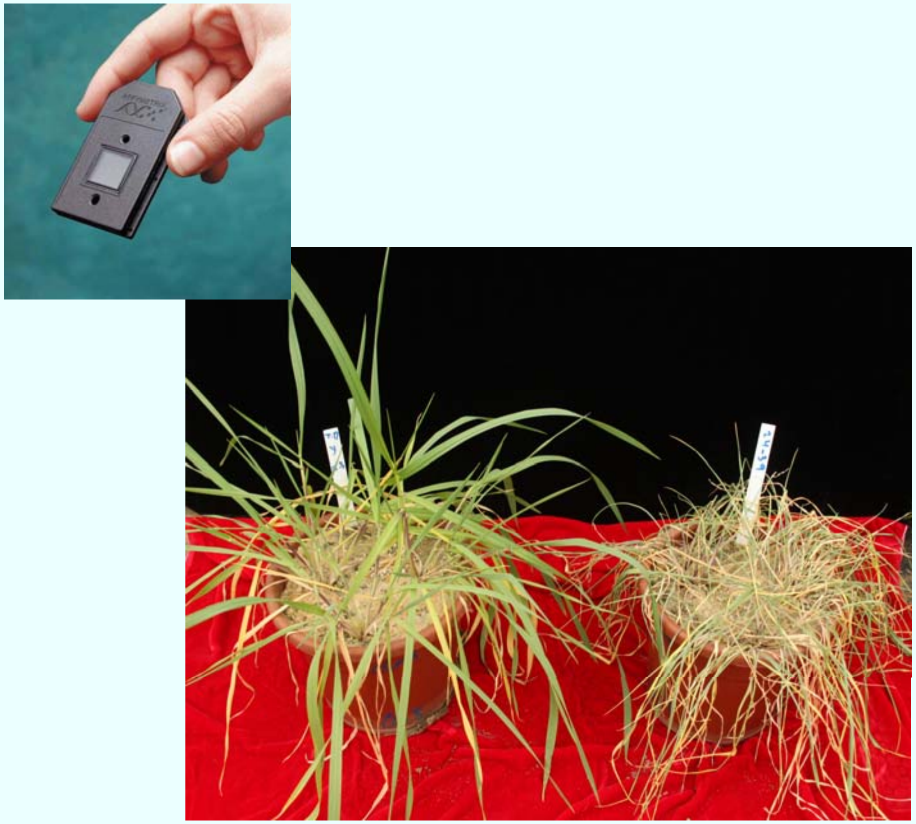
Drought stress is a major challenge for several fields in plant biology. In this study we used a multidisciplinary approach:

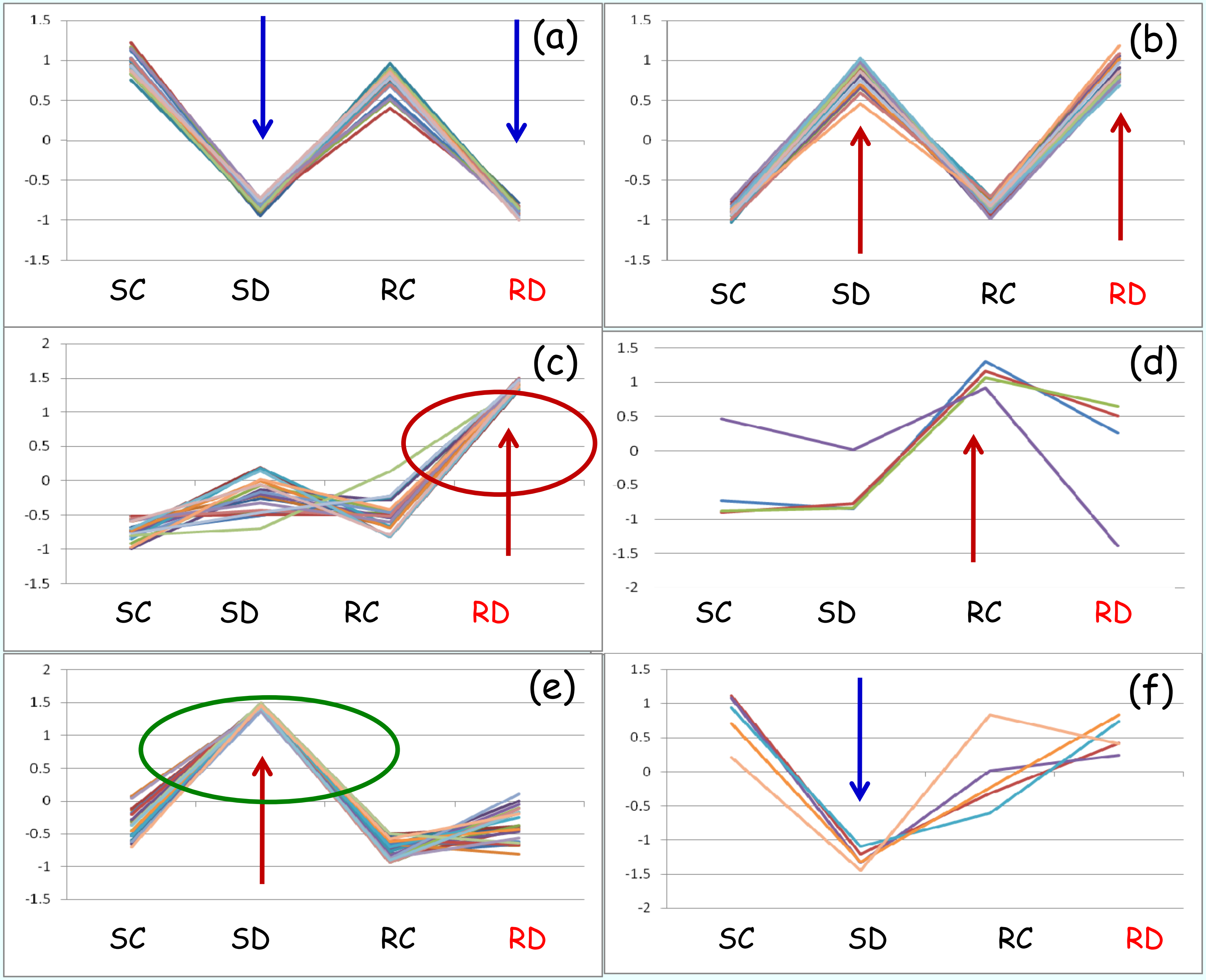
☐ Population genetics
☐ QTL mapping
☐ Functional genomics

Assisted by

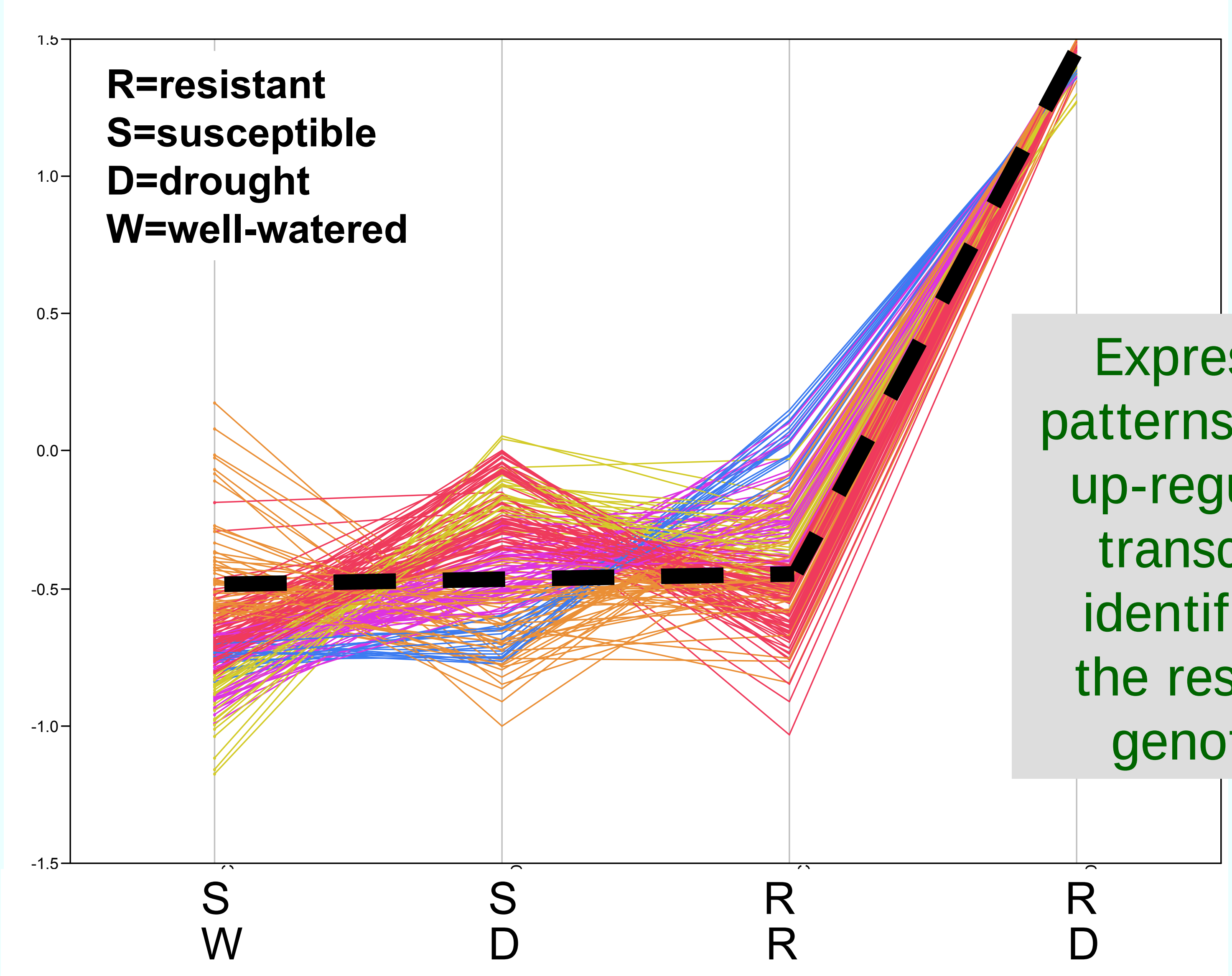
☐ Plant physiology

☐ Metabolomics




(a)
1.5
1
0.5
0
-0.5
-1
-1.5
SC
SD
RC
RD
(b)
1.5
1
0.5
0
-0.5
-1
-1.5
SC
SD
RC
RD
(c)
2
1.5
1
0.5
0
-0.5
-1
-1.5
SC
SD
RC
RD
(d)
1.5
1
0.5
0
-0.5
-1
-1.5
-2
SC
SD
RC
RD
(e)
2
1.5
1
0.5
0
-0.5
-1
-1.5
SC
SD
RC
RD
(f)
1.5
1
0.5
0
-0.5
-1
-1.5
-2
SC
SD
RC
RD

Transcriptome analysis to reveal candidate genes for drought stress tolerance in emmer wheat

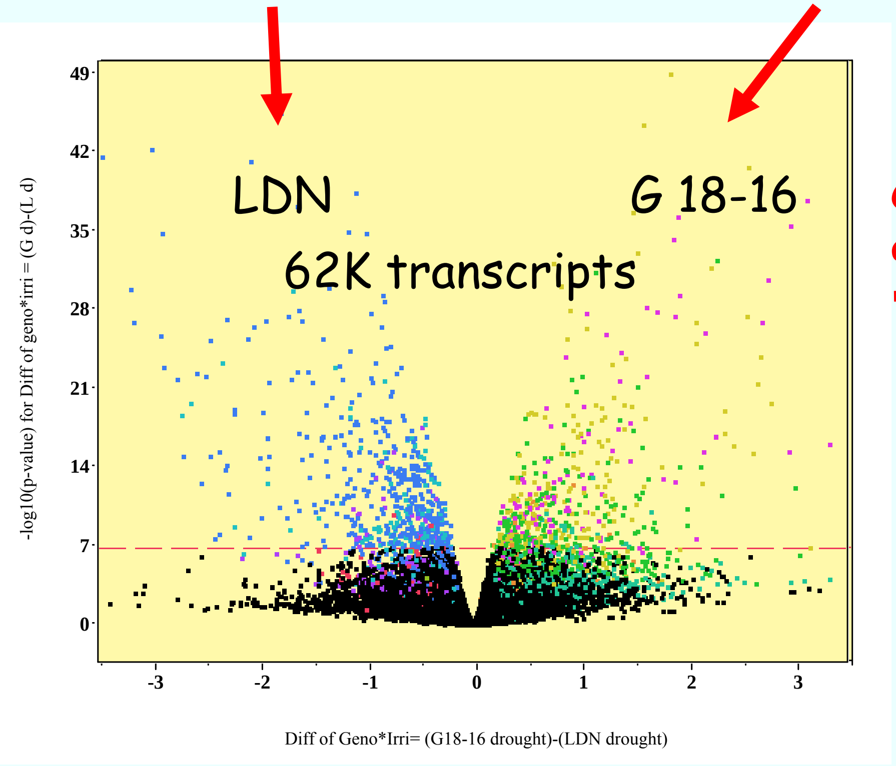


Expression patterns of 300 up-regulated transcripts identified in the resistant genotype

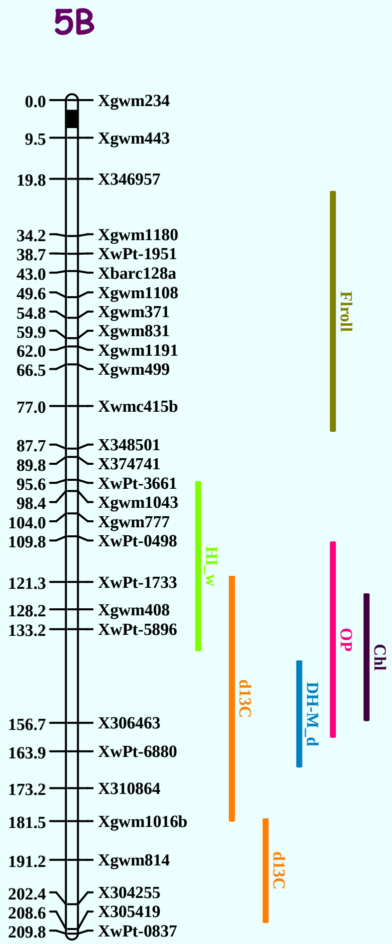
Transcriptome Analysis of Response to Drought

Langdon

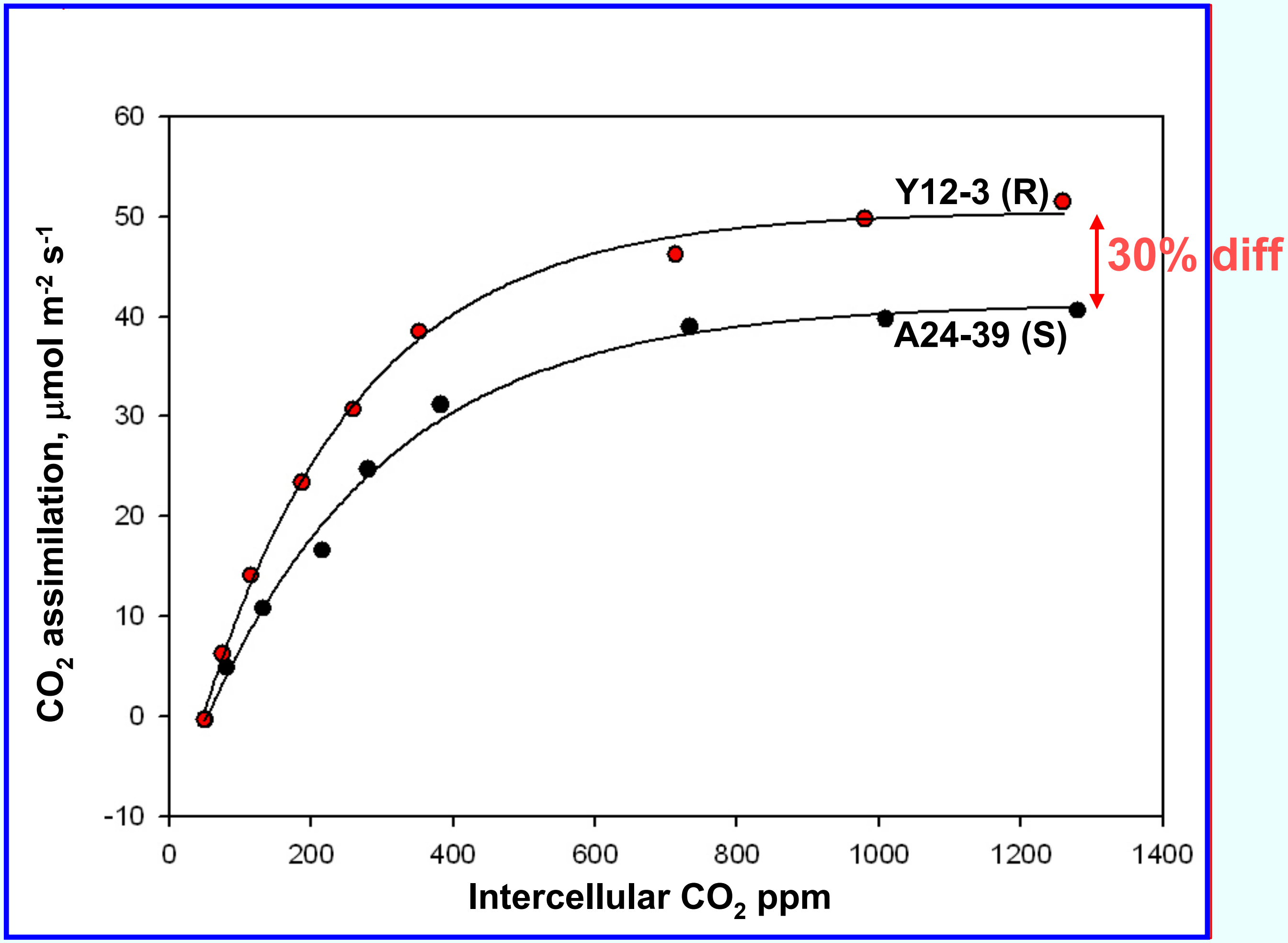
Gitit 18-16



Candidate gene mapping and QTL co-localization



Photosynthetic rate of drought resistant vs. drought susceptible genotypes under drought stress



GO enrichment analysis

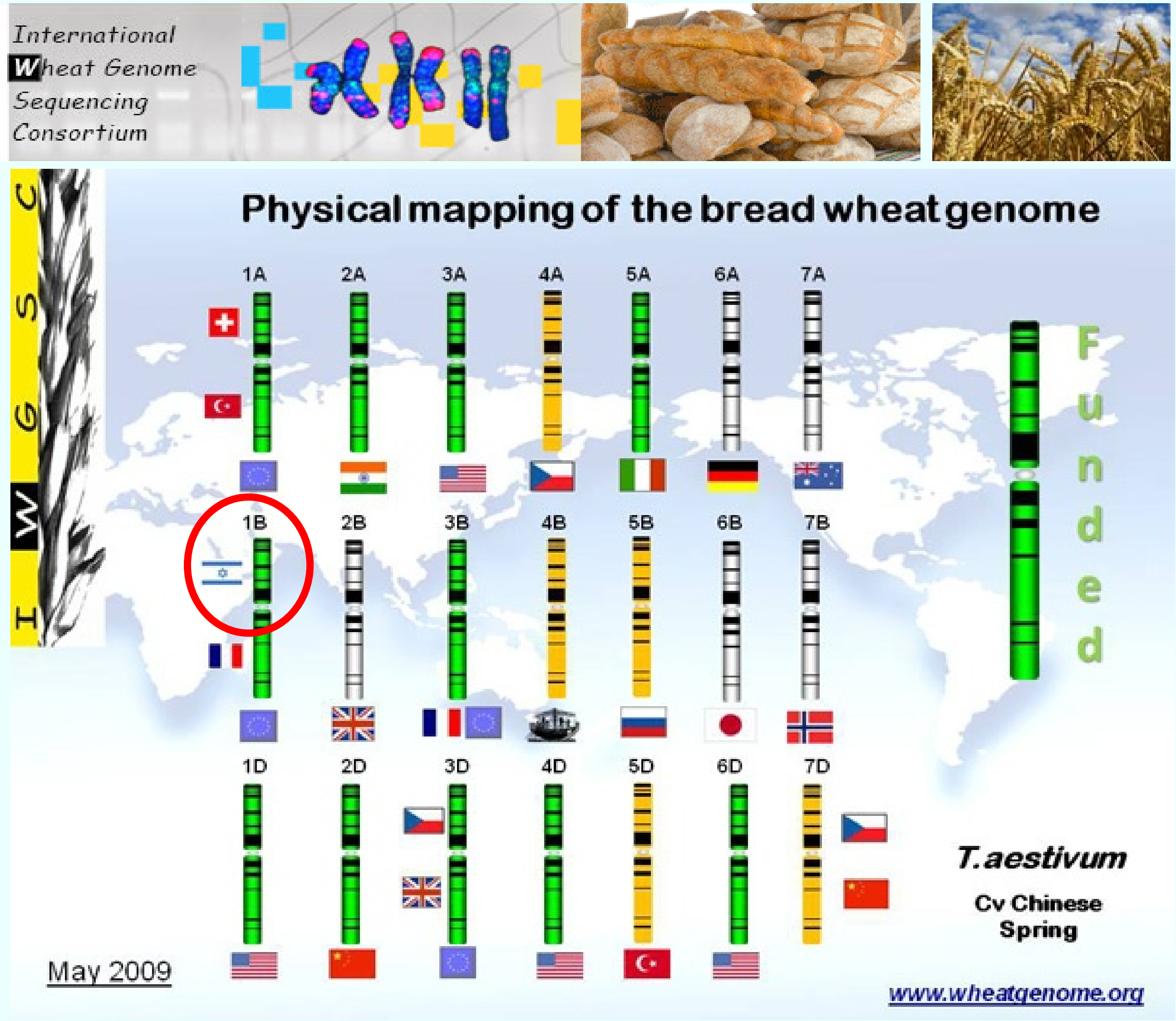
Biological process GO term	**Number of transcripts and Cluster frequencies**		**Genome frequency of 56,985 used rice genes**
	Out of 837 DETs - R	**Out of 514 DETs – S**	
Response to stress	104, 12.4%	61, 11.9%	2516, 4.4%
Response to endogenous stimulus	**76, 9.1%**	**n.s.**	**3963, 7.0%**
Response to abiotic stimulus	**70, 8.4%**	**48, 9.3%**	**2308, 4.1%**
Response to biotic stimulus	47, 5.6%	33, 6.4%	2019, 3.5%
Response to external stimulus	27, 3.2%	14, 2.7%	797, 1.4%
Response to stimulus	180, 21%	121, 23.5%	6182, 10.8%

Adaptive mechanisms of drought response identified in the resistant genotype

- ✓ Abscisic acid (ABA) dependent stomatal closure
- ✓ Cell wall adjustment
- ✓ Cuticular wax formation
- ✓ Lignification
- ✓ Osmoregulation and cell homeostasis
- ✓ Energy transfer
- ✓ Dehydration protection
- ✓ Maintained metabolism and catalytic activity
- ✓ Delayed senescence

International effort for wheat genome physical mapping

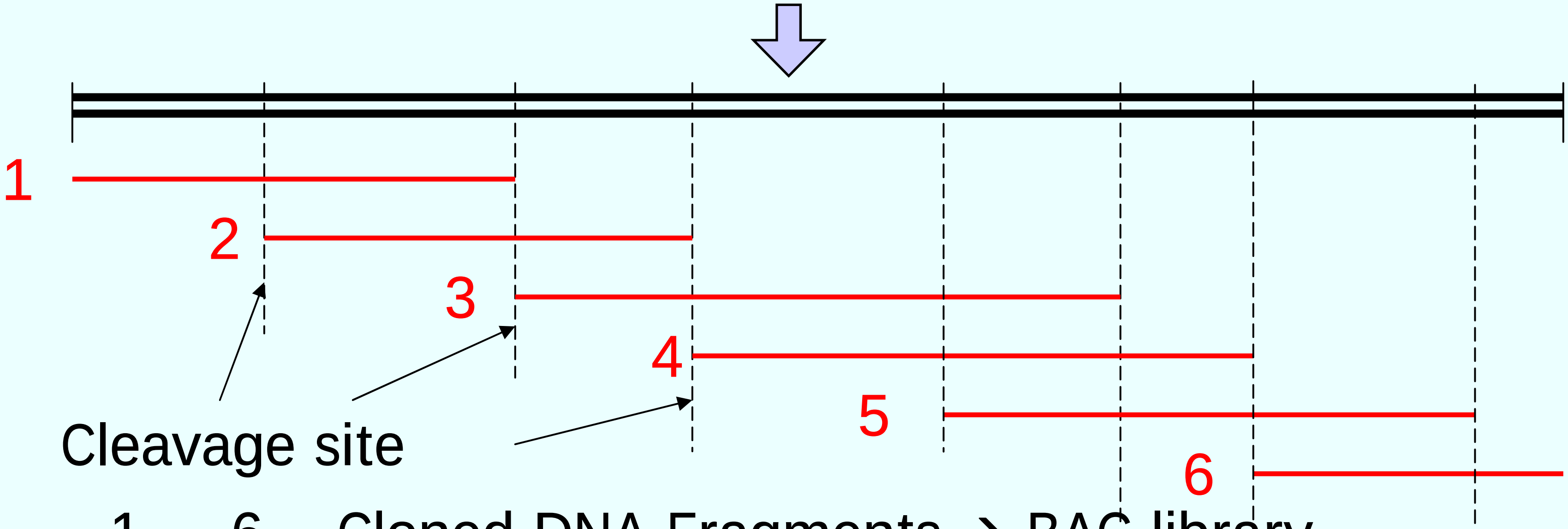
Chromosome 1BS - University of Haifa



Physical mapping with BAC clones

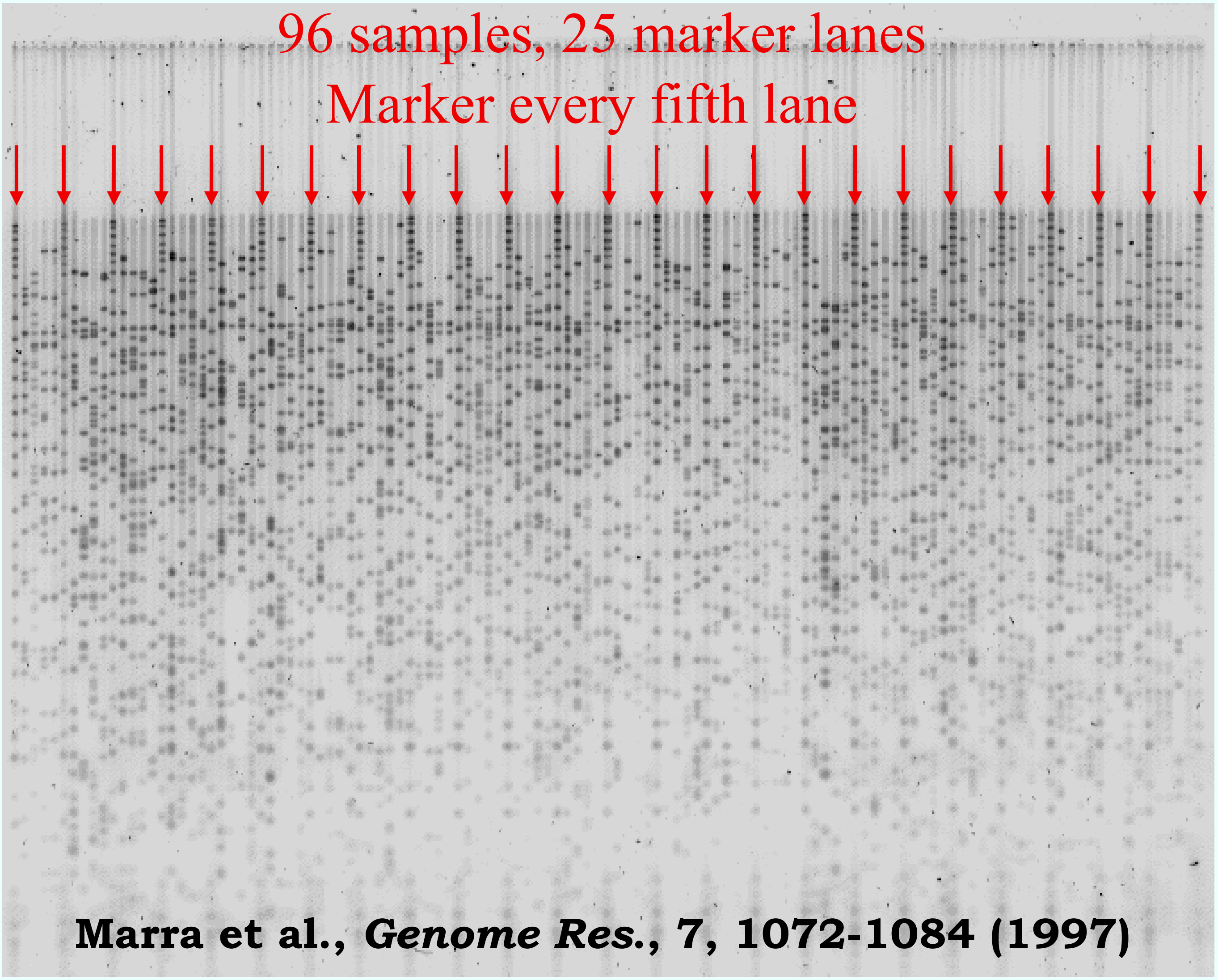
Genomic library: Breaking the DNA, cloning the fragments in BACs, and ordering

Let us cut the isolated DNA with a restriction enzyme taken at a low concentration many sites will remain unrestricted



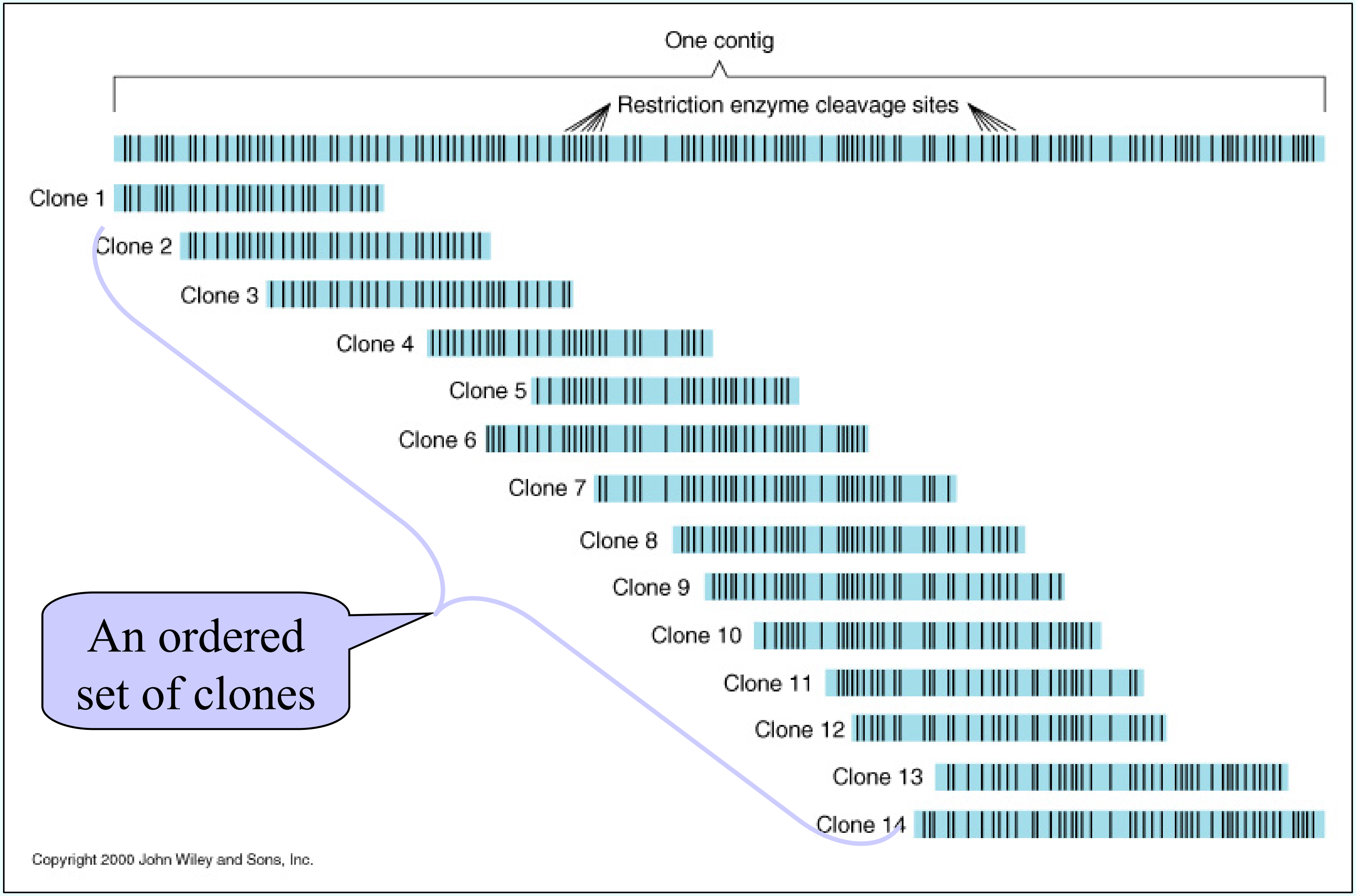
1,…,6 Cloned DNA Fragments → BAC library

BAC Fingerprinting by Fragment Separation



Marra et al., *Genome Res.*, 7, 1072-1084 (1997)

Building a contig map

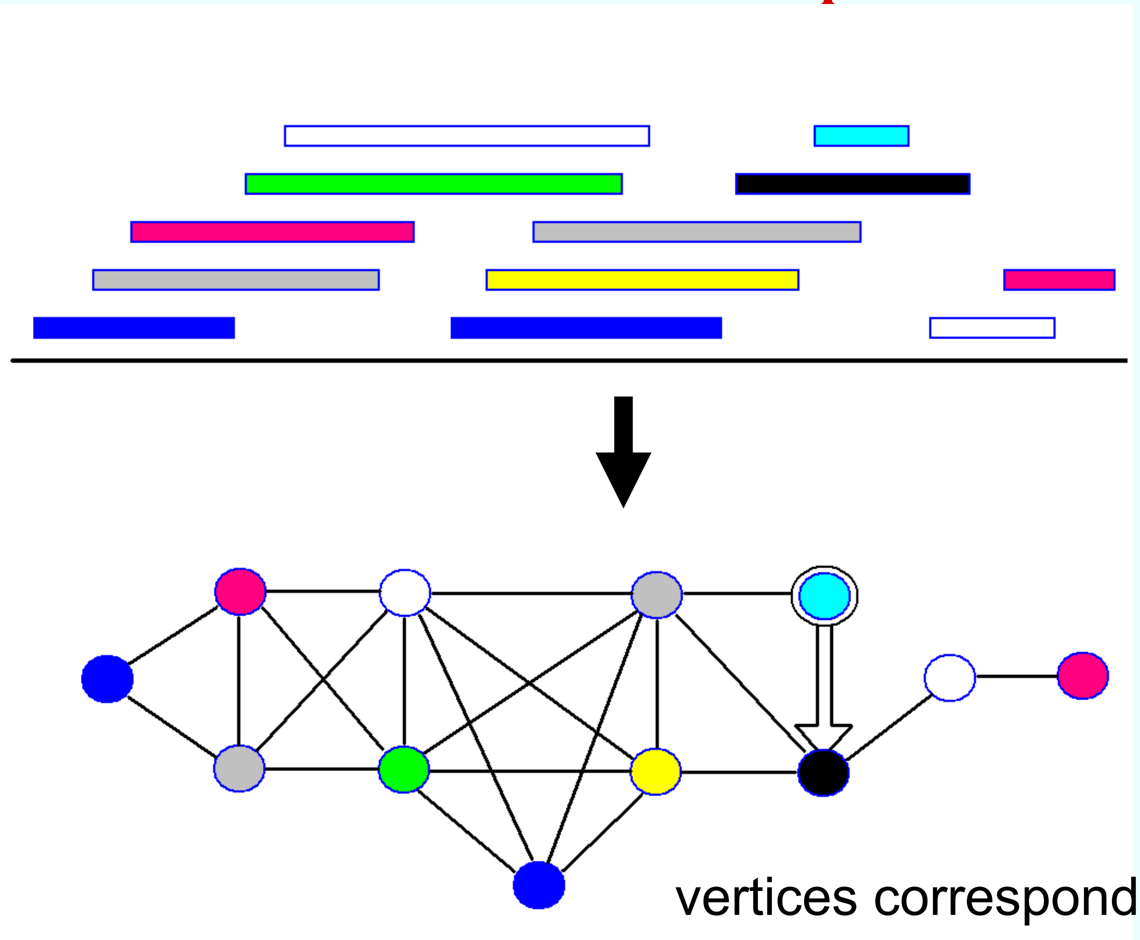


Genome mapping problems are computationally challenging

> *"… We have been looking at the assemblies of large genomes … and **for every 'draft' genome** we look at, we find hundreds - and sometimes thousands - **of mis-assemblies**".*

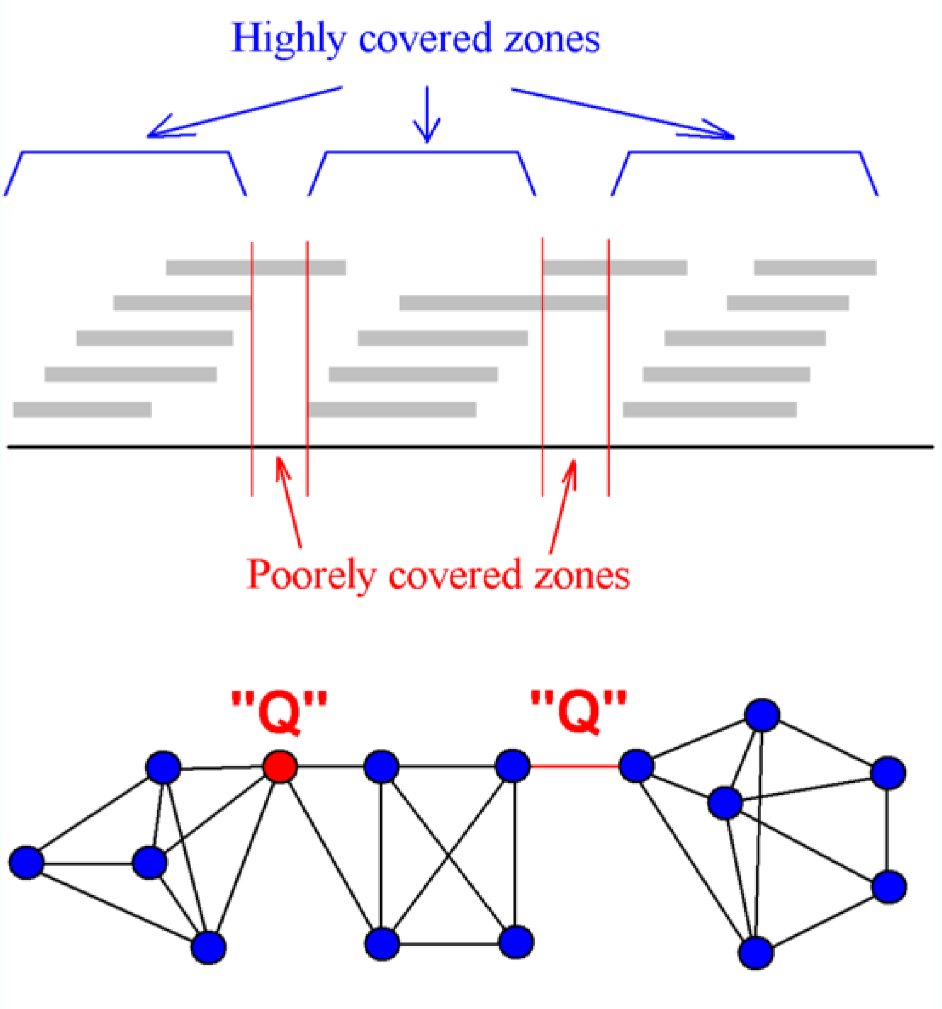
Salzberg & Yorke (2005) Beware of mis-assembled genomes. *Bioinformatics*, **21**: 4320-4322

Network representation of significant clone overlaps



vertices correspond to clones and edges – to significant clone overlaps

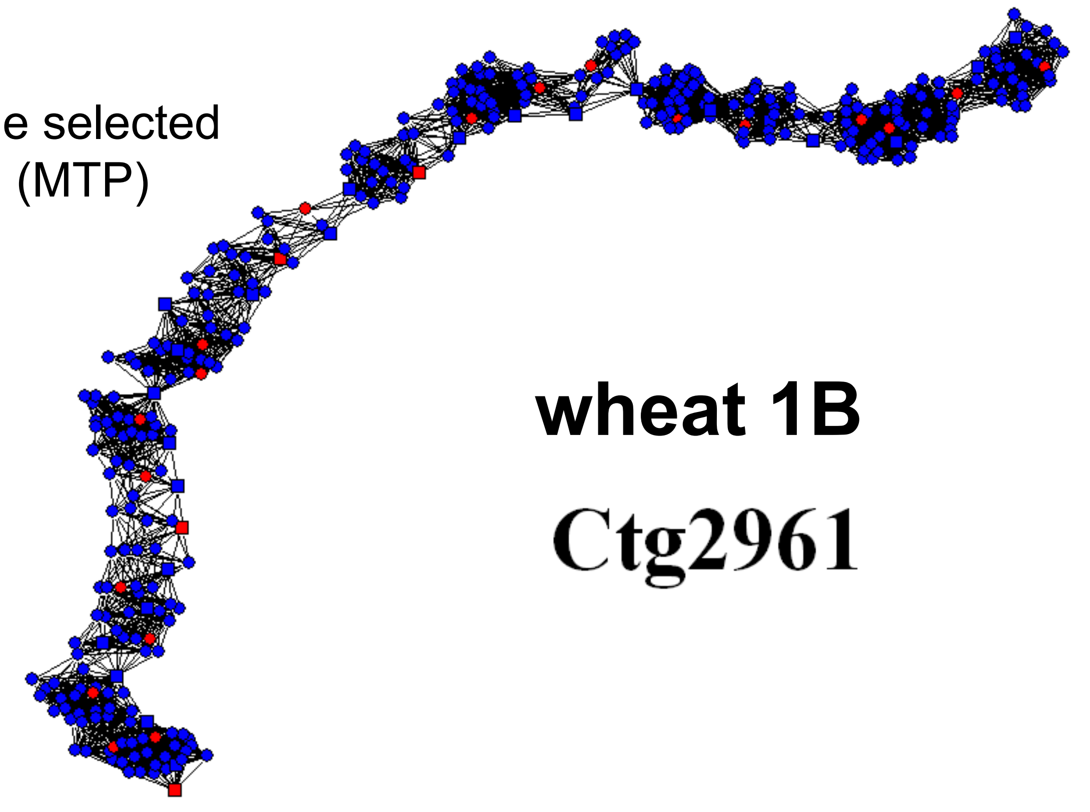
Putative Q-clones and Q-overlaps



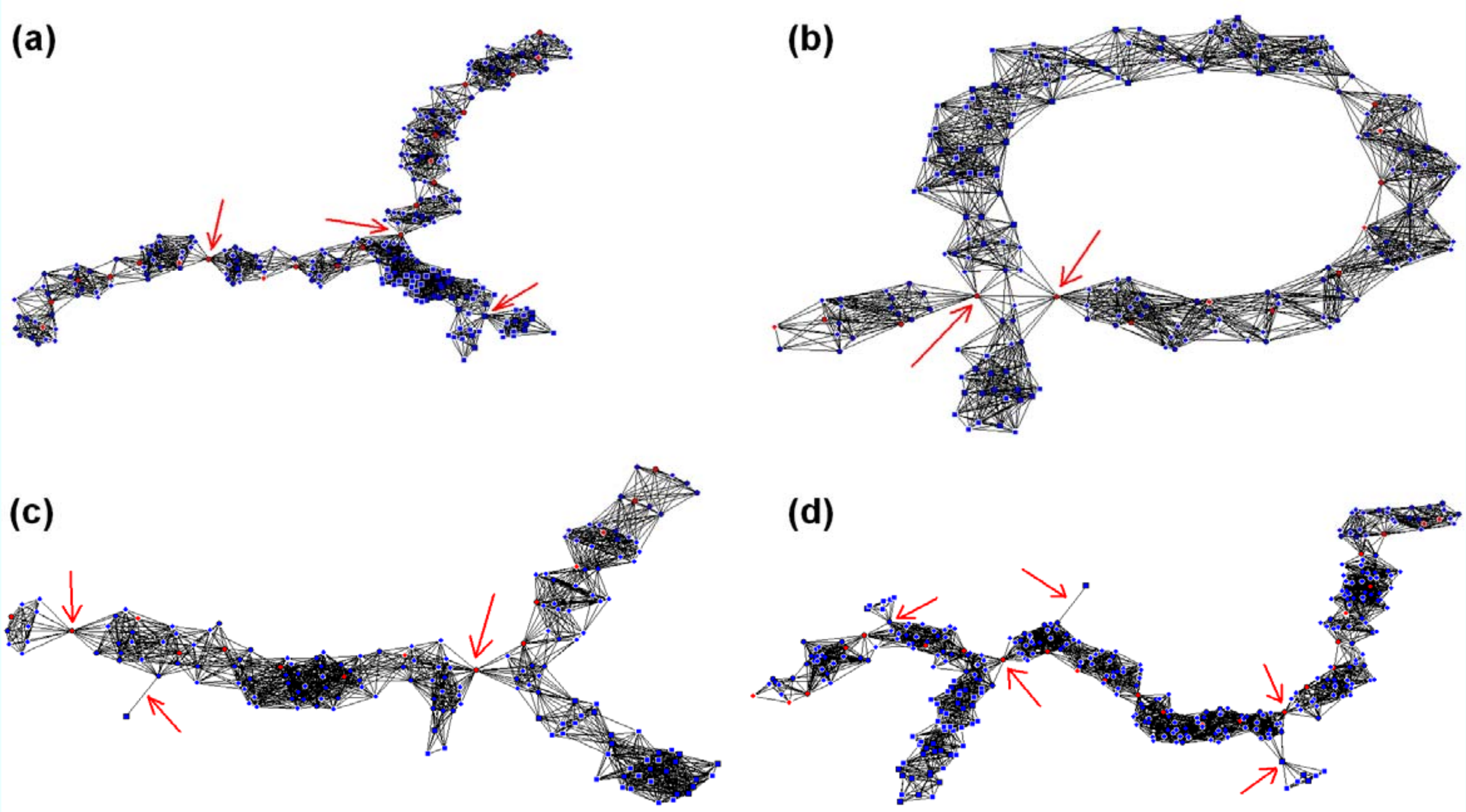
Frenkel et al. 2010

Network representation of significant clone overlaps

- 🔵 clones
- 🔴 clones from the selected diametric path (MTP)



Identification of contig non-linearity



אוניברסיטת חיפה
University of Haifa

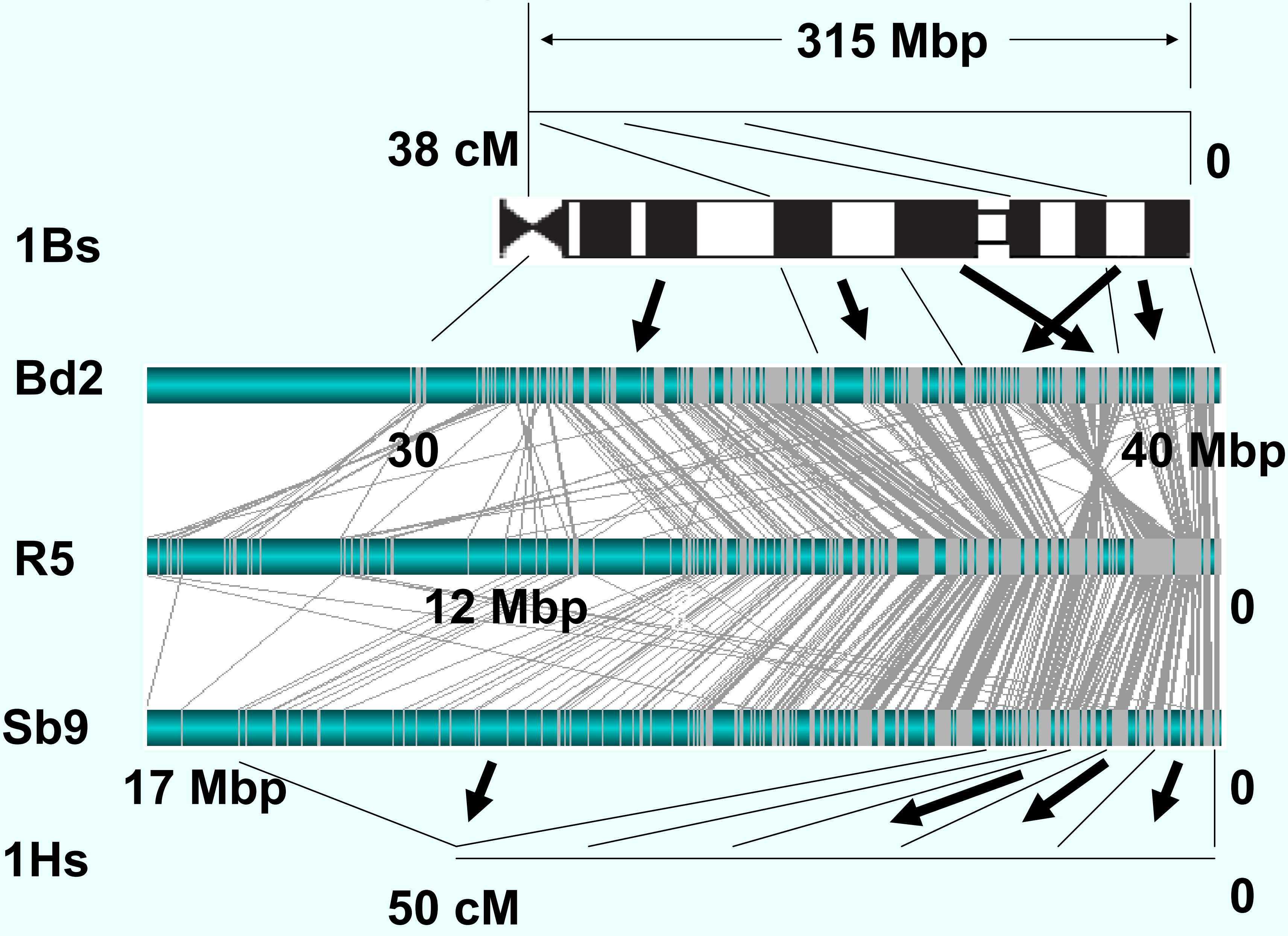
1BS assembly: FPC vs. LTC

In total 49,412 clones (av. size 113 kb), coverage 17.7

	FPC	**LTC**
Contigs with ≥6 clones	517	385
Clones in contigs	33,262	33,912
Mean clones/contig	64.3	88.1
Clones in MTPs	3,647	3,827
Coverage by MTP	270 Mb (86%)	283 Mb (90%)

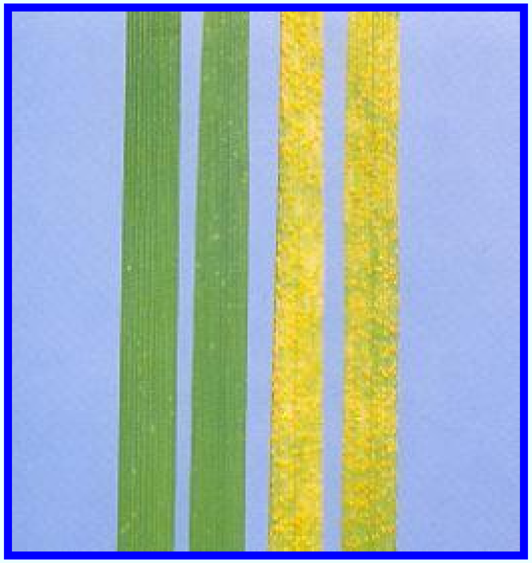
MTPs were constructed by LTC using e-25 cutoff

Parallel genomes for 1BS



Genomic Organization and Diversity of Stripe Rust Resistance Genes Derived from Wild Emmer Wheat

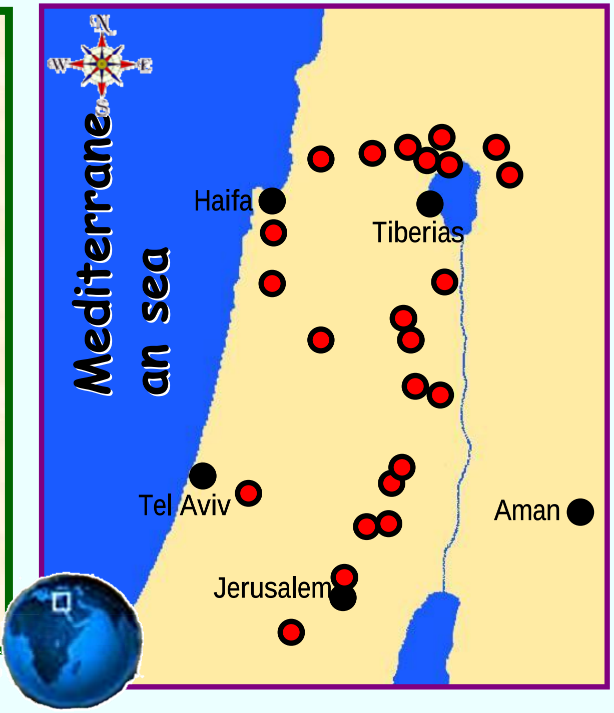
How many genes are involved?
What is their genomic organization?



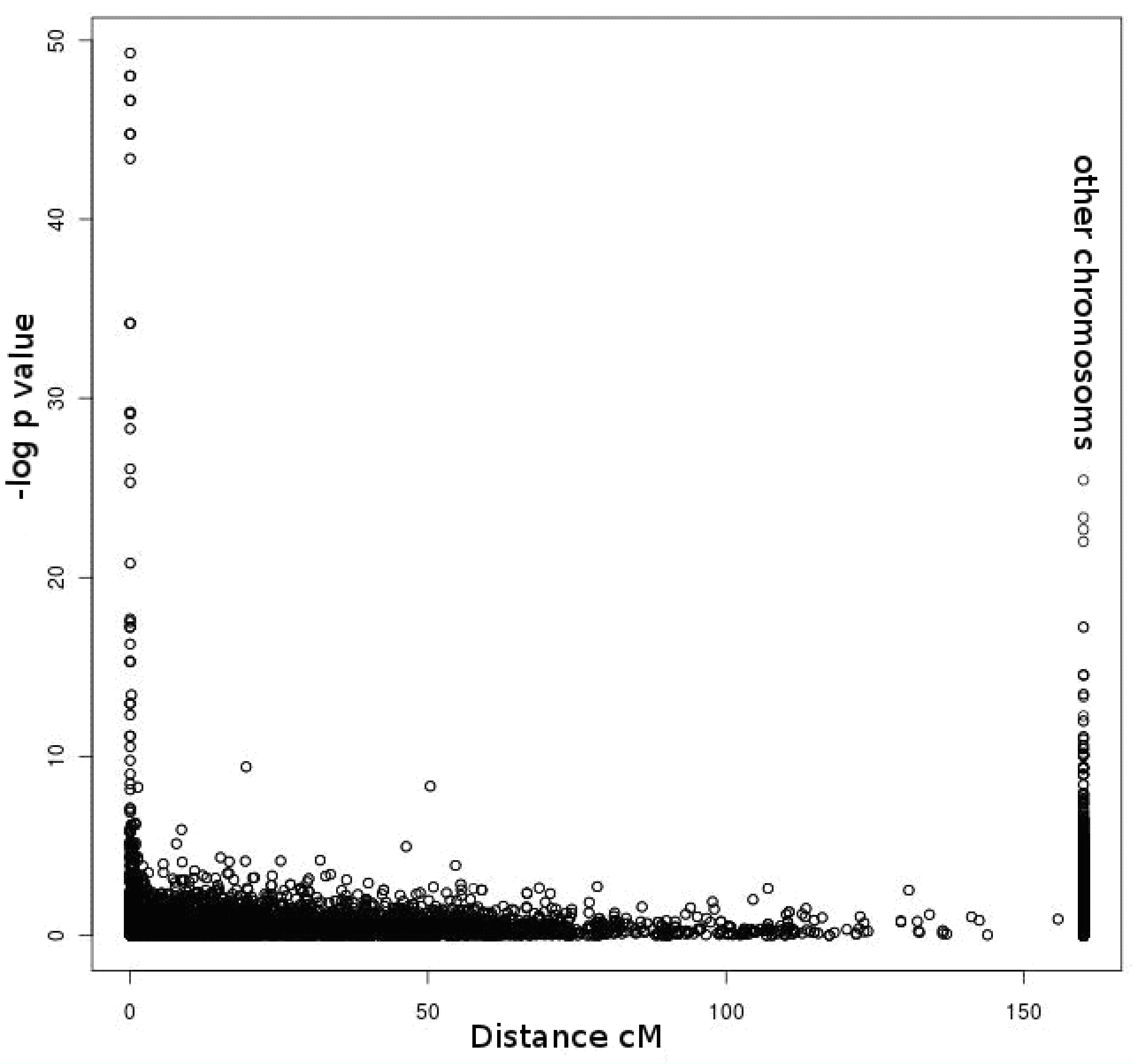
Theoretical calculations

=

11 novel Yr genes

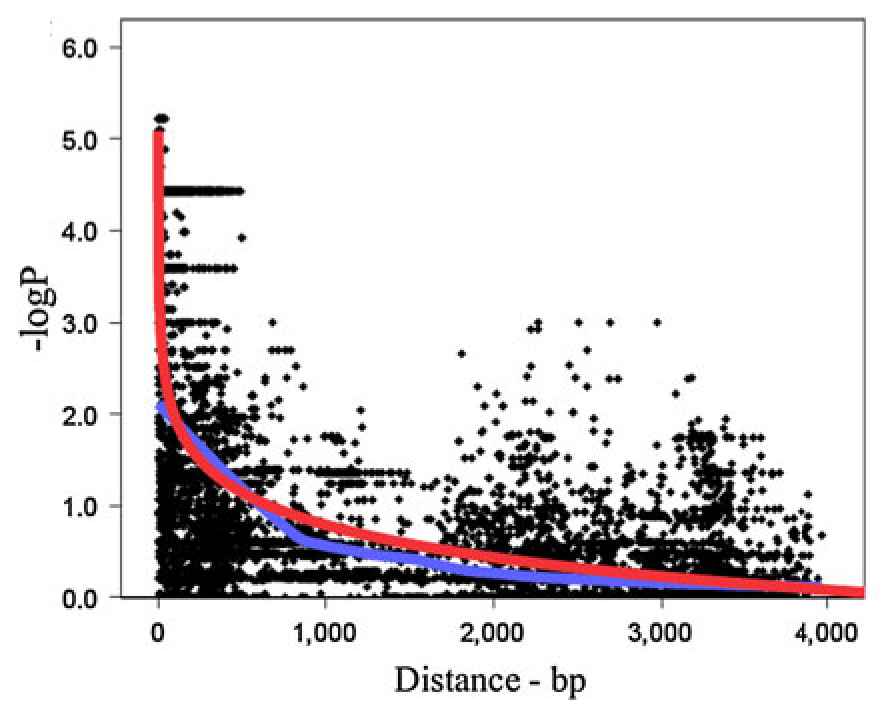


Prospects for association mapping in *T. dicoccoides*

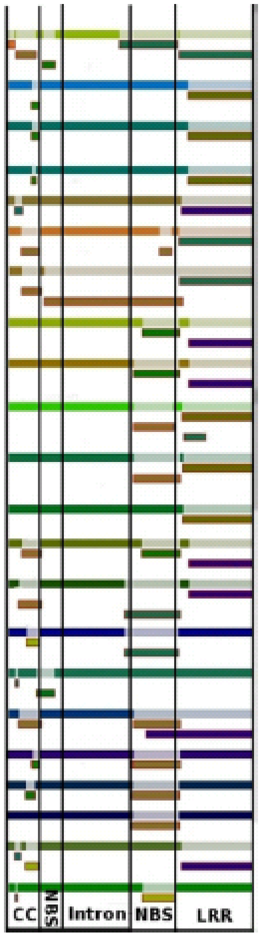


Decay in LD
between SNP
markers
across genome

Prospects for association mapping in *T. dicoccoides*



Decay in LD within leaf rust resistance gene LR10 in Israeli populations of wild emmer



Sela et al. TAG 122:175–187, 2011

Participating groups

Tzion Fahima, Tamar Krugman, Abraham Korol, Eviatar Nevo, Vladimir Frenkel, Yuchun Li, Junhua Peng, Olga Raskina, Alexander Belyaev, Hanan Sela, Assaf Distelfeld, Dina Raats, **Institute of Evolution, University of Haifa, Israel**

Jorge Dubcovsky, Cristobal Uauy, Lynn Epstein, **Department of Plant Science, University of Calofornia, Davis**

Yehoshua Saranga, Zvi Peleg, Asaf Avneri, **Institute of Plant Science & Genetics, Hebrew University of Jerusalem, Israel**

Boulos Chalhoub, Veronique Chague, Sandrine Balzergue, Nathalie Boudet, **Organization and Evolution of Plant Genomes, Unité de Recherche en Génomique Végétale, Evry, France**

Etienne Paux, Catherine Feuillet**, INRA, Amélioration et Santé des Plantes, Clermont-Ferrand, France**

Aaron Fait, Lydia Quansah, **Ben Gurion University of the Negev, Israel**

Marion Röder, **Institute of Plant Genetics & in Crop Plant Research (IPK), Gatersleben, Germany**

Благодарю за терпение

Please, visit our website: http://evolution.haifa.ac.il