QUALITY WHEAT
CRC PROJECT REPORT

Workshop:
Molecular Technologies for the Wheat Industry

19th February 1999

Quality Wheat CRC & The University of Sydney

Date: February 1999

QWCRC Report No: 22
Copy No: 58

CONFIDENTIAL
(Not to be copied)

Quality Wheat CRC has taken all reasonable care in preparing this publication. Quality Wheat CRC expressly disclaims all and any liability to any person for any damage, loss or injury (including economic loss) arising from their use of, or reliance on, the contents of this publication.

Quality Wheat CRC Ltd.
Professional Development Program
Workshop: Molecular Technologies for the Wheat Industry
Friday February 19, 1999

Program:

Time	Session	Speaker
9.30-10.20	Introduction to Molecular Genetics	*(Clare Johnson)*
10.20-10.40	tea	
10.40-11.30	Origins and relatives of wheat, and conventional breeding strategies	*(Fred Stoddard)*
11.30-12.20	Molecular tags, and design for rapid breeding	*(Peter Sharp)*
12.20-1.00	lunch	
1.00-1.50	Bioinformatics - Gene discovery in the area of wheat flour quality	*(Rudi Appels)*
1.50- 2.40	Proteomics, protein technologies, potential and limitations	*(Brad Walsh)*
2.40-3.00	tea	
3.00-3.50	Application of molecular technologies in diagnostics etc for the wheat industry	*(Thomas Giersch)*
3.50-5.00	Reception	

Hope you enjoy the day!
Clare Johnson
Education & Training Coordinator

The University of Sydney

Molecular Genetics Overview

Clare Johnson

Introduction to:

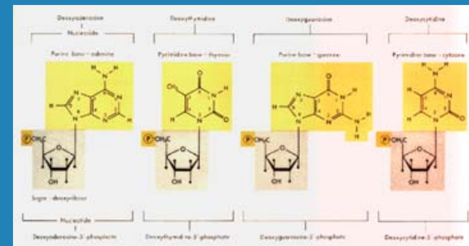
- DNA structure
- Sense direction
- Sequencing and PCR
- Restriction enzyme specificity / mapping
- Coding / expression
- Genome / Proteome
- Concepts of homology - linear, motif, structural

Nucleic Acid Bases

- 4 bases
- Plus deoxyribose

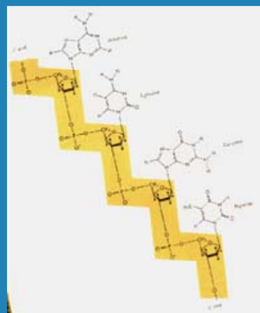
Adenine ➡ Deoxyadenosine 5′ phosphate
Thymine ➡ Deoxythymidine 5′ phosphate
Cytosine ➡ Deoxycytidine 5′ phosphate
Guanine ➡ Deoxyguanidine 5′ phosphate

Deoxyribonucleotides



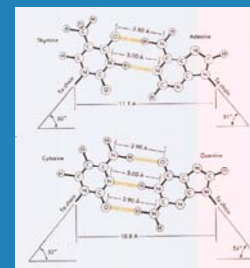
Common: sugar phosphate backbone

- Phosphodiester bonds link 5′ ⟶ 3′



Base-pairing

- A pairs with T by 2 H bonds
- G pairs with C by 3 H bonds



Complementary strands and direction

- one strand implies the other

Sense strand

5′ AAGACTACGTCGGATAGATCCCA 3′
3′ TTCTGATGCAGCCTATCTAGGGT 5′

Antisense strand

- strands go in opposite directions

Strand extension

5′ → 3′

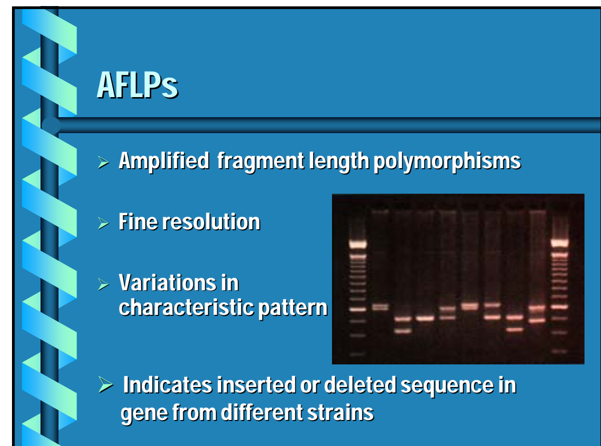
AAGACTACGTCGGATAGATCCCAGTAAGT
TTCTGATGCAGCCTATCTAGGGTCATTCA

PCR: geometric progression

AAGACTACGTCGGATAGATCCCAGTAAGT
melt apart (94oC)
TTCTGATGCAGCCTATCTAGGGTCATTCA

5′ TTCTGATGC → 3′
AAGACTACGTCGGATAGATCCCAGTAAGT

TTCTGATGCAGCCTATCTAGGGTCATTCA
3′ ← CCAGTAAGT 5′



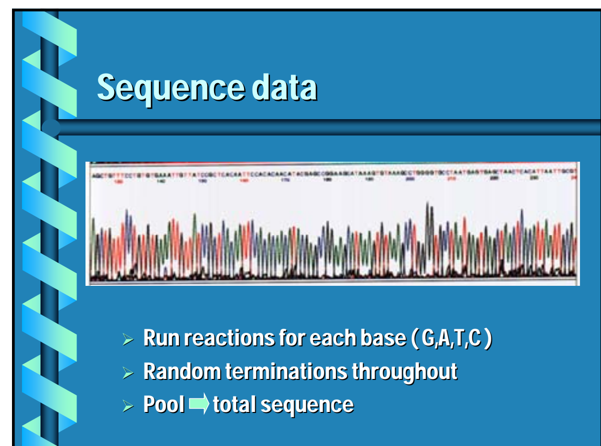
Sequencing

5′ → 3′

AAGACTACG
TTCTGATGCAGCCTATCTAGGGTCATTCA

AAGACTACGTCGGATAG*
TTCTGATGCAGCCTATCTAGGGTCATTCA

AAGACTACGTCGGATAGATCCC*
TTCTGATGCAGCCTATCTAGGGTCATTCA



DNA motifs: Restriction enzyme specificity

mirror image

TACGTCGGATCCATCCCAGT
ATGCAGCCTAGGTAGGGTCA

Restriction enzyme specificity

BamHI

TACGTCGGATCCATCCCAGT
ATGCAGCCTAGGTAGGGTCA

TACGTCG GATCCATCCCAGT
ATGCAGCCTAG GTAGGGTCA

Complementary sticky ends

Flexibility

Bam HI

TACGTCGGATCCATCCCAGT
ATGCAGCCTAGGTAGGGTCA

Bgl II

TACGTCAGATCTATCCCAGT
ATGCAGTCTAGATAGGGTCA

Restriction Mapping

- 6 base palindromes occur rarely (1 in 4^6 =4.1kb)
- 8 base palindromes more rarely (1 in 4^8=65.6kb)
- Digests ⇒ characteristic pattern of fragments

RFLPs

- Restriction fragment length polymorphisms
- Variations in characteristic pattern
- Indicates inserted or deleted sequence in gene from different strains

Coding and expression

AACAAATGGTCGGATATATCCCAGTAAGT
TTGTTTACCAGCCTATATAGGGTCATTCA

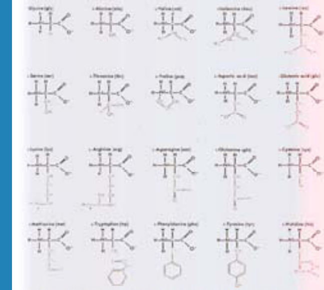
AACAAATG.GTC.GGA.TAT.ATC.CCA.GTA.AGT
Met Val Gly Tyr Ile Pro Val Ser

Total 20 common amino acids plus STOP from 64 possible codons

Genetic code

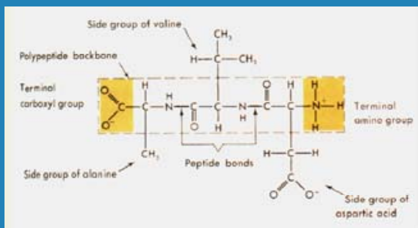
First Position (5′end)		Second	Position		3rd Position (3′ end)
	U (T)	C	A	G	
U (T)	Phe Phe Leu Leu	Ser Ser Ser Ser	Tyr Tyr STOP STOP	Cys Cys STOP Trp	U (T) C A G
C	Leu Leu Leu Leu	Pro Pro Pro Pro	His His Gln Gln	Arg Arg Arg Arg	U (T) C A G
A	I le I le I le Met	Thr Thr Thr Thr	Asn Asn Lys Lys	Ser Ser Arg Arg	U (T) C A G
G	Val Val Val Val	Ala Ala Ala Ala	Asp Asp Glu Glu	Gly Gly Gly Gly	U (T) C A G

Amino Acids



Amino Acid to Protein

- Coding direction is N ➡ C



Genome / Proteome

Different genes expressed in different

- tissues
- developmental stages
- metabolic states

Concepts of homology

Similarity may be

- global (linear, whole cDNA)

Protein a
Protein b

- local (funtional motifs within cDNA)

Protein c
Protein d

Concepts of homology

Similarity may be

- structural (including active triad sites)

- far fewer structures than sequences

Wheat evolution and breeding

Fred Stoddard
The University of Sydney
and
Quality Wheat CRC Ltd

Wheat is inbreeding

- Fixed genotypes
- Cultivars contain identical plants
  - within crops
  - between seasons
- Breeding program must be designed accordingly

Some really basic genetics

- Higher organisms are diploid (2n)
- one set of chromosomes (n) from each parent
- at fertilization, n + n --> 2n
- Genes are in pairs

Meiosis or reduction-division

- chromosomes pair and exchange genetic information
- one member of each pair goes to each pole
- thus each gamete is unique

Paired genes

- Often "dominant" and "recessive"
  - one dose is enough
  - AA = Aa aa
- others "additive"
  - AA > Aa > aa

Multigenic traits

Hexaploidy

- *T. aestivum* is a "segmental hexaploid"
  - (diploid three times)
- This cubes the probability of finding a random mutation:
  1 in 10^6 per genome = 1 in 10^{18}
  = 30 billion tonnes of wheat

Example: waxy starch

- Barley, rice, maize: waxy is wxwx
- Wxwx and WxWx are normal
- So in wheat we want
  $wx_A wx_A wx_B wx_B wx_D wx_D$ (aabbdd)
  but we can't distinguish the starch of AABBdd from AAbbDD or aaBBDD
- We need extra tools for wheat

Conventional breeding strategy

- Many basic requirements
  - Appropriate phenology (vrn, ppd)
  - Disease resistances
  - Grain colour, milling yield
  - Hardness, glutenin composition for end use
- New quality attributes

Typical wheat program

- Cross to generate variability
- Inbreed to fix new combinations
- Select for appropriate types
- Multiply seed for release
- Roseworthy as an example
  - Accelerated - 10 years instead of 13

MOLECULAR TAGS & RAPID BREEDING

TAGS:

	RFLP	RAPD	SSR	AFLP
Assay	RE + hybridization	PCR with random 10mers	PCR with specific primers	RE + PCR with random selective bases
Type of Polymorphism	Single base, insertions & deletions	Single base, insertions & deletions	Repeat length	Single base, insertions & deletions
Level of Polymorphism	medium	medium	high	low
Inheritance	Co-dominant	Dominant	Co-dominant	Dominant
DNA required	10ug	10ng	10ng	1ug
DNA sequence required?	No	No	Yes	No
Radioactive detection?	Yes	No	No	Yes/no

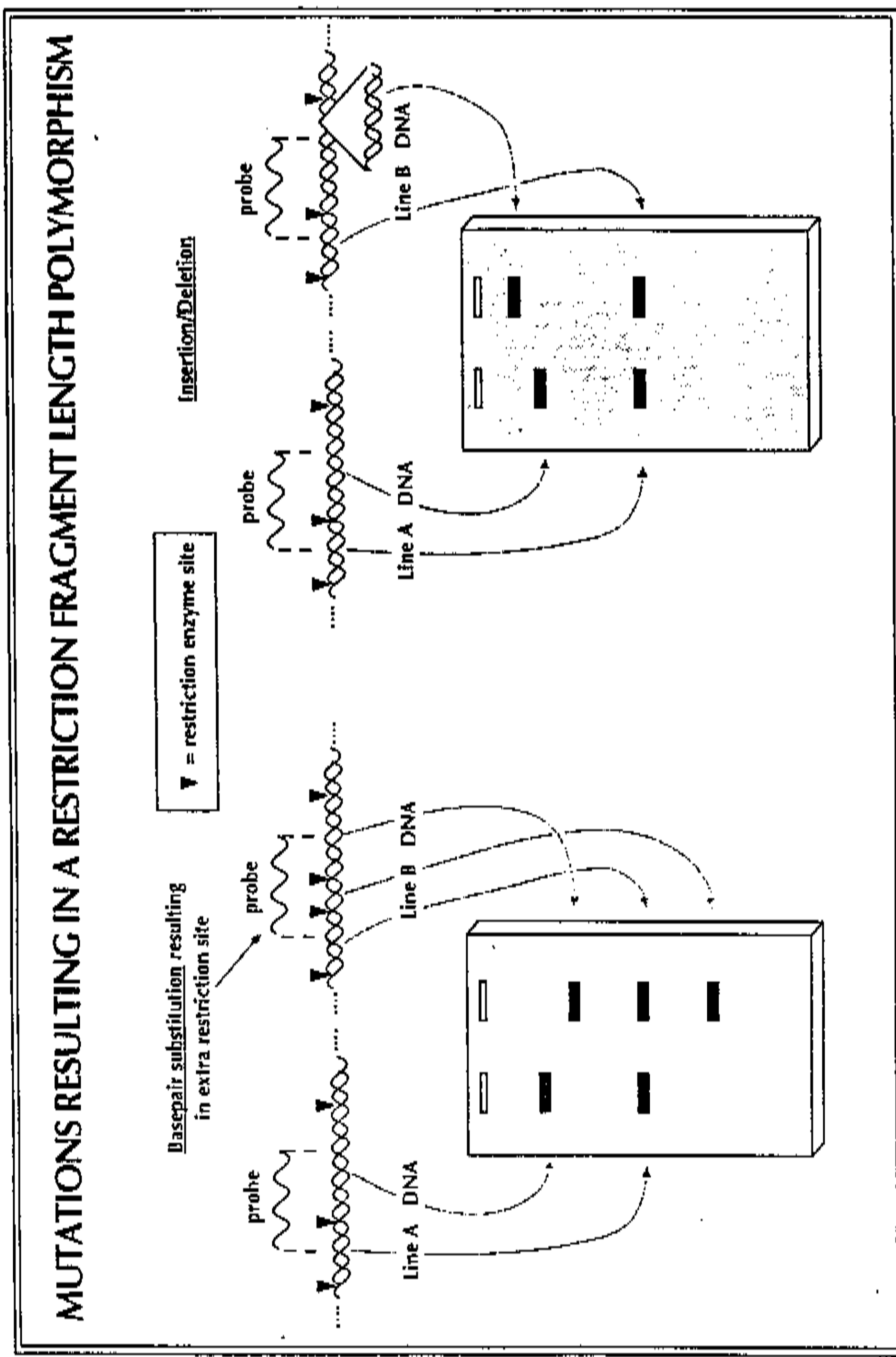


Fig. 12. Molecular origin of RFLPs

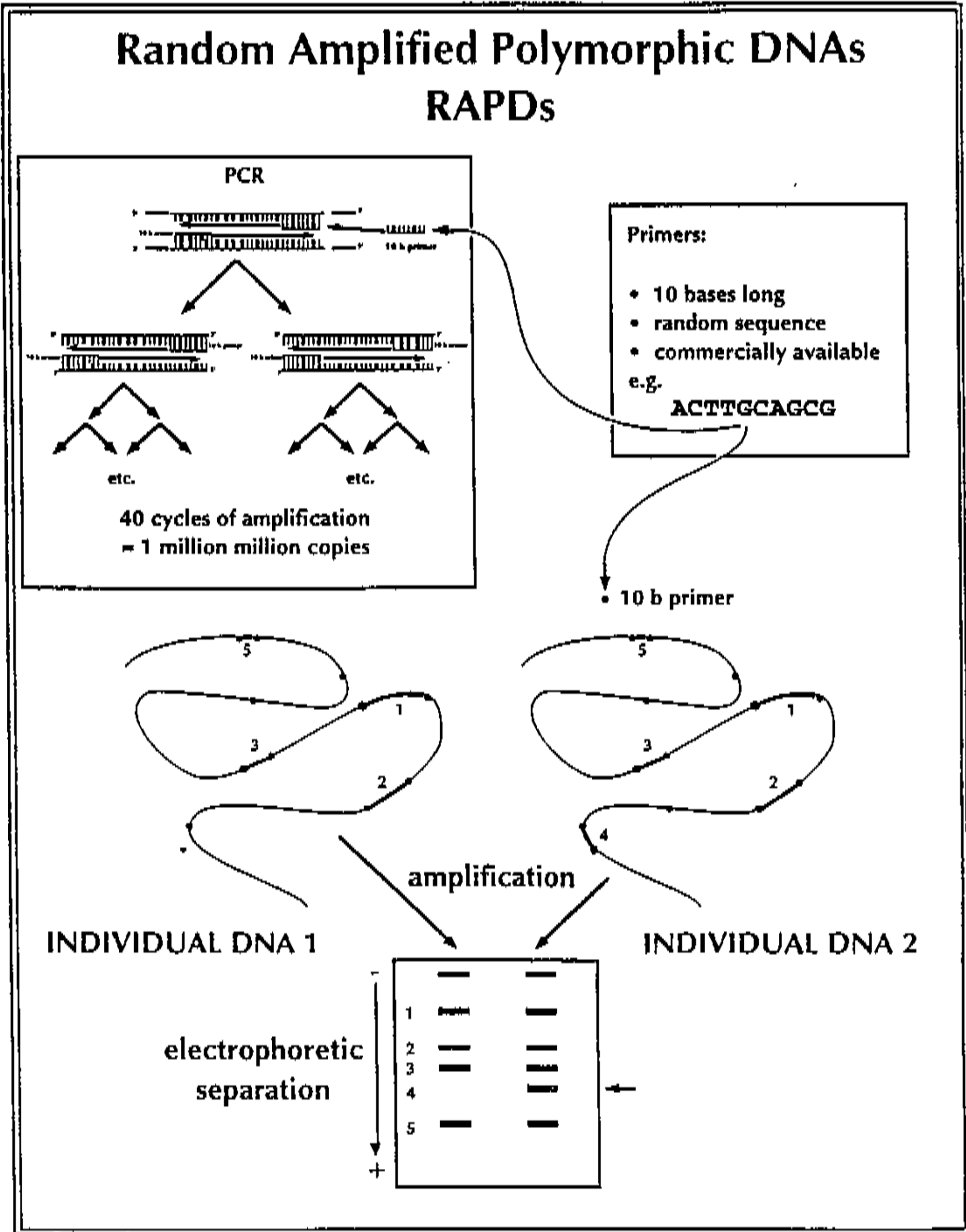
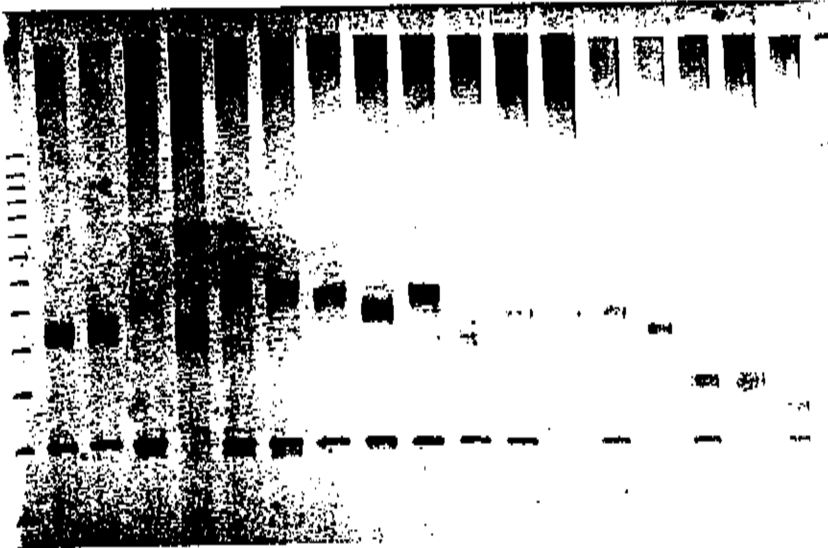


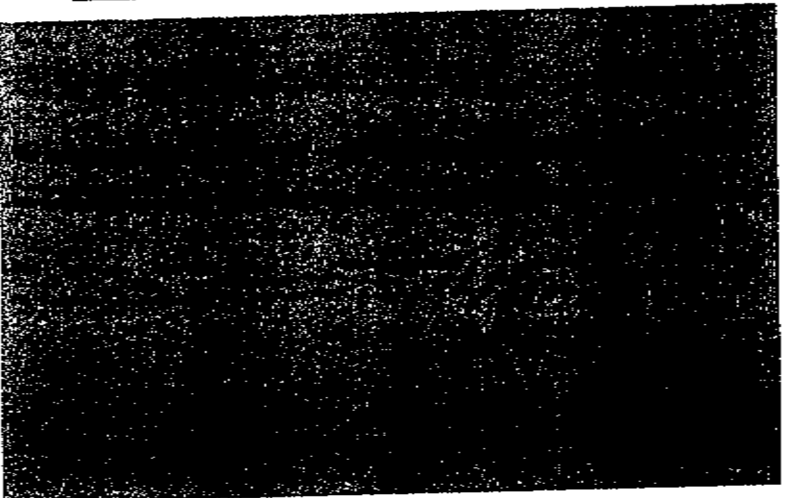
Fig. 1



PLBR 348 FIG 1

AFLP segregation in Halberd X Stiletto

H S ______ F3 bulks ______



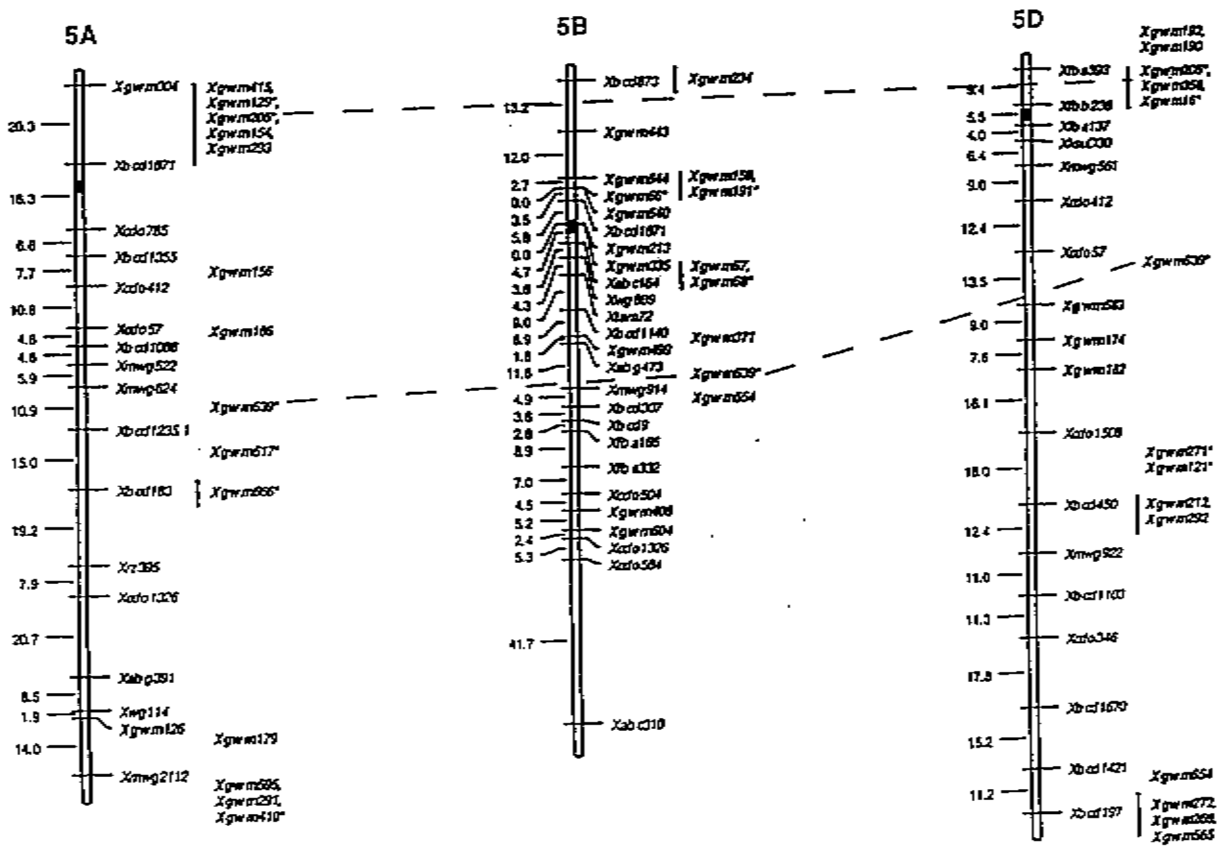


FIGURE 1.—*Continued.*

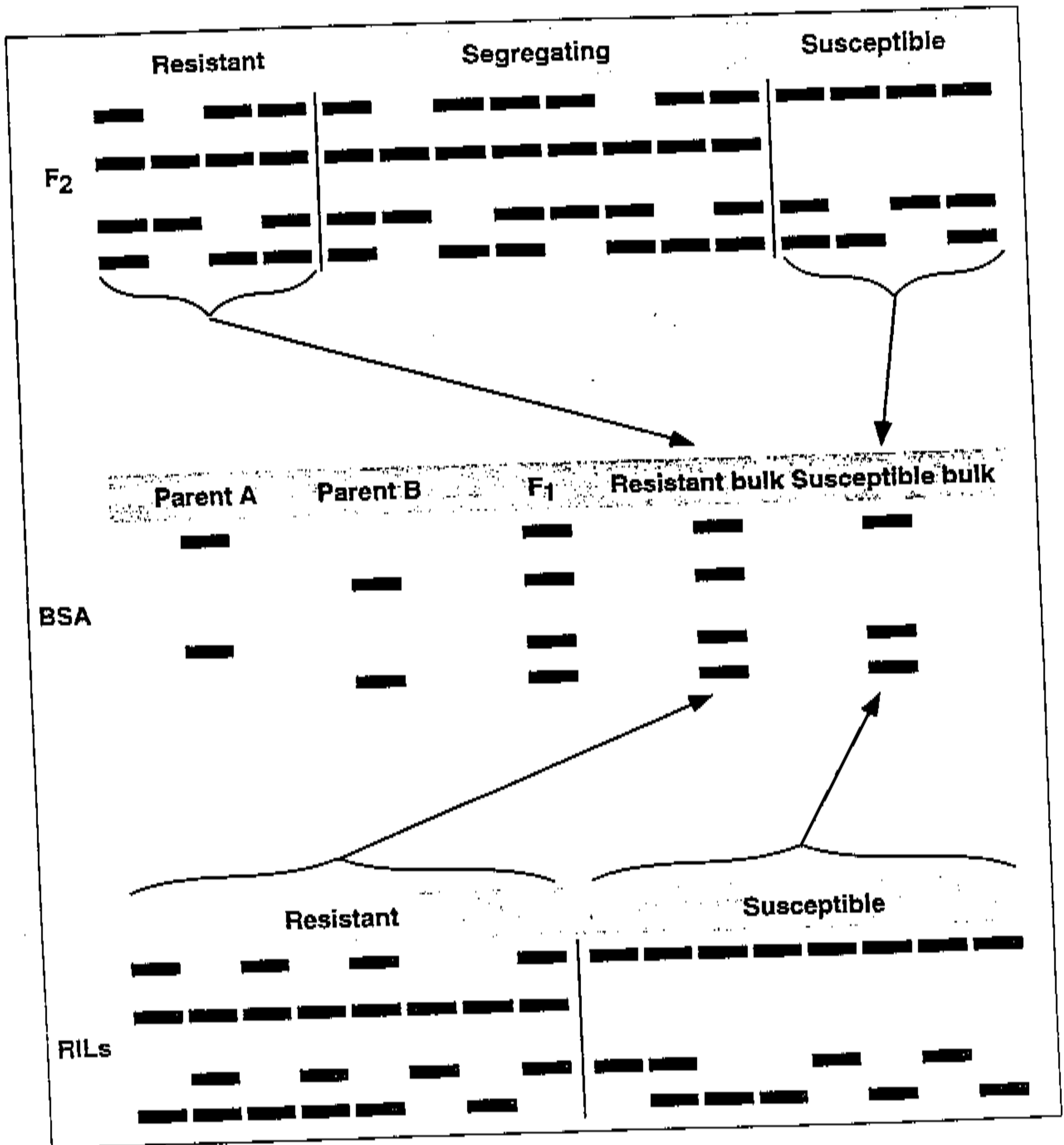


Fig. 2. *Bulk Segregant Analysis.*

USES OF "TAGS" IN SELECTION:

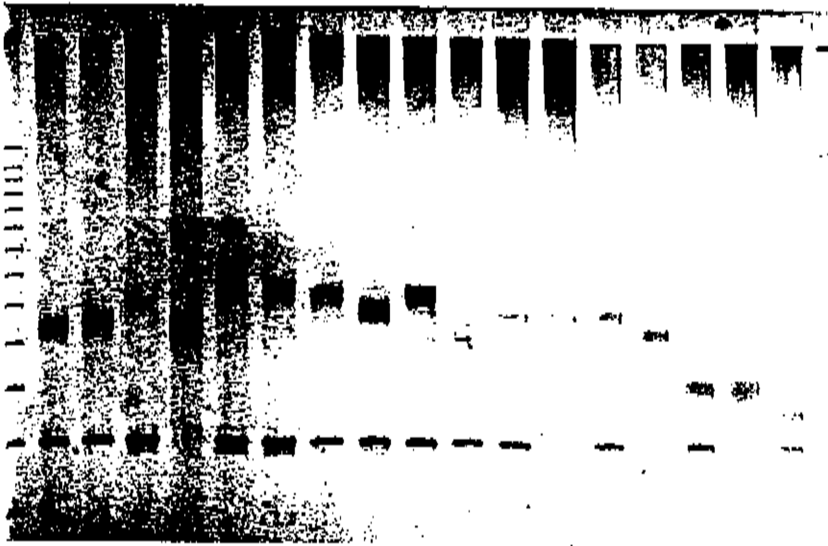
OVERALL AIM IS TO INCREASE CERTAINTY

- CONFIRMATION OF HYBRIDITY

- +VE SELECTION

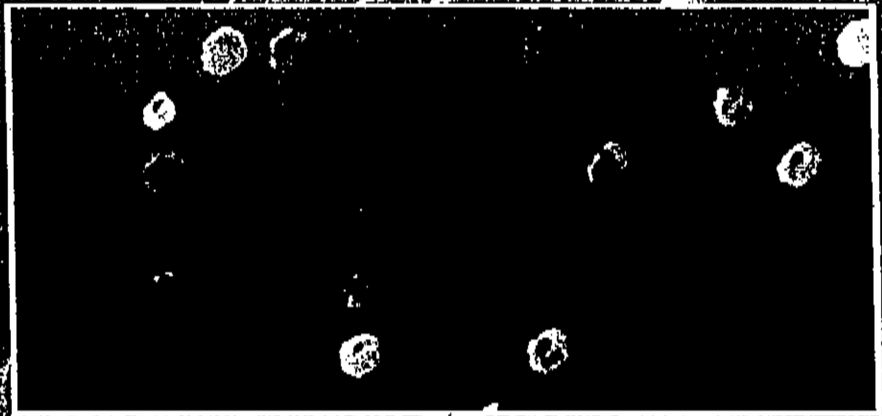
- – VE SELECTION (BACKGROUND "CLEAN-UP", ESPECIALLY IN BACKCROSSING)

6, 18



PLBR 348 FIG 1

EXECUTIVE'S SPECIAL PROJECTS FUNDS PROPOSAL



Genomi

Gene overy

Aus

Gene discovery in the area of wheat quality

- Large-scale sequencing of expressed genes

Wheat endosperm cDNA library from 8/10/12 days post-anthesis.

25,000 clones grided into 384 well microtiter plates for long term storage and large-scale sequencing

Sequences from 1000 clones are close to completion

Gene discovery in the area of wheat quality

- Large-scale sequencing of expressed genes

File from sequencer is sent to ANGIS and vector sequences "clipped"

Files are filtered for quality and the ones that pass are subjected to a BLAST X and BLAST N search

The BLAST X search translates a DNA sequence to protein sequences (3 reading frames, given the 5' end of the cDNA is defined) and then uses these to screen the world data base

```
>sp|P08079|GDBO_WHEAT GAMMA-GLIADIN PRECURSOR >pir||PS0094 gamma-gliadin     722
           precursor (clone pW10) - wheat (fragment) >gi|170736 (M16060)
           gamma-gliadin [Triticum aestivum]
           Length = 251

  Plus Strand HSPs:

 Score = 66 (30.2 bits), Expect = 3.8e-10, Sum P(3) = 3.8e-10
 Identities = 16/40 (40%), Positives = 19/40 (47%), Frame = +1

Query:    28 MKTMFXLALLAFTATSAVAQLYTTCSQGYGQCQQQPQPQP 147
             MKT+  L +LA   T   A +     S         QQQP PQP
Sbjct:     1 MKTLLILTILAMAITIGTANMQVDPSSQVQWPQQQPVPQP 40

 Score = 46 (21.1 bits), Expect = 3.8e-10, Sum P(3) = 3.8e-10
 Identities = 10/29 (34%), Positives = 14/29 (48%), Frame = +1

Query:   130 QPQPQPXMNTCAAFLXQCIQTPYVQSQMW 216
             QP  Q  +N C  FL Q  +    + S +W
Sbjct:   146 QPSLQQVNPCKNFLLQQCXPVSLVSSLW 174

 Score = 82 (37.6 bits), Expect = 3.8e-10, Sum P(3) = 3.8e-10
 Identities = 14/27 (51%), Positives = 19/27 (70%), Frame = +2

Query:   221 QASCQLMXXQCCQPLAXIXEQARCQAV 301
             Q+ CQ+M  QCCQ LA I +Q +C A+
Sbjct:   180 QSDCQVMRQQCCQQLAQIPQQLQCAAI 206
```

```
>sp|P06659|GDBB_WHEAT GAMMA-GLIADIN B PRECURSOR >pir||EEWTG gamma-gliadin B     972
           precursor - wheat >gi|170708 (M13713) gamma-gliadin B precursor
           [Triticum aestivum]
           Length = 291

  Plus Strand HSPs:

 Score = 116 (53.1 bits), Expect = 1.5e-46, Sum P(3) = 1.5e-46
 Identities = 24/28 (85%), Positives = 24/28 (85%), Frame = +1

Query:     4 LLIQTILVMAITIATANMQVDPSGQVPW 87
             LLI TIL MAITIATANMQ DPSGQV W
Sbjct:     4 LLILTILAMAITIATANMQADPSGQVQW 31

 Score = 239 (109.5 bits), Expect = 1.5e-46, Sum P(3) = 1.5e-46
 Identities = 44/46 (95%), Positives = 45/46 (97%), Frame = +2

Query:   451 LNPCKNFLLQQCKPVSLVSSLWSNILPRSDCQVMRQQCCQQLAQIP 588
             LNPCKNFLLQQCKPVSLVSSLWS+ILP SDCQVMRQQCCQQLAQIP
Sbjct:   146 LNPCKNFLLQQCKPVSLVSSLWSIILPPSDCQVMRQQCCQQLAQIP 191

 Score = 125 (57.3 bits), Expect = 1.5e-46, Sum P(3) = 1.5e-46
 Identities = 25/32 (78%), Positives = 28/32 (87%), Frame = +3

Query:   594 LXCATIHSIVHSIIMRKEQQEQRQGVQXLVPL 689
             L CA IHS+VHSIIM++EQQEQ QGVQ LVPL
Sbjct:   194 LQCAAIHSVVHSIIMQQEQQEQLQGVQILVPL 225
```

```
>sp|P52589|PDI_WHEAT PROTEIN DISULFIDE ISOMERASE PRECURSOR (PDI) /     746
           DOLICHYL-DIPHOSPHOOLIGOSACCHARIDE-PROTEIN GLYCOTRANSFERASE
           (GLYCOSYLATION SITE-BINDING CHAIN) (GSBP) >gi|508975 (U11496)
           protein disulfide isomerase [Triticum aestivum] >prf||2106410A
           protein disulfide isomerase [Triticum aestivum]
           Length = 515

  Plus Strand HSPs:

 Score = 549 (251.5 bits), Expect = 3.8e-89, Sum P(3) = 3.8e-89
 Identities = 107/109 (98%), Positives = 107/109 (98%), Frame = +2

Query:     2 REAEGIVEYLKKQVGPASKEIKAPEDATYLEDGKIHIVGVFTEFSGTEFTNFLELAEKLR 181
             REAEGIVEYLKKQVGPASKEIKAPEDATYLEDGKIHIVGVFTEFSGTEFTNFLELAEKLR
Sbjct:   136 REAEGIVEYLKKQVGPASKEIKAPEDATYLEDGKIHIVGVFTEFSGTEFTNFLELAEKLR 195

Query:   182 SDYDFGHTVHANHLIHGDAAXERPLVRLXKPFDELVVDSKDFDVSALEK 328
             SDYDFGHTVHANHLPRGDAA ERPLVRL KPFDELVVDSKDFDVSALEK
Sbjct:   196 SDYDFGHTVHANHLPRGDAAVERPLVRLFKPFDELVVDSKDFDVSALEK 244

 Score = 185 (84.7 bits), Expect = 3.8e-89, Sum P(3) = 3.8e-89
 Identities = 35/42 (83%), Positives = 35/42 (83%), Frame = +3

Query:   327 KFIDASXTPXXVTFDXNPDNHPYLLKXFQSNXPXAMLFXNFS 452
             KFIDAS TPX VTFD NPDNHPYLLK FQSN P AMLF NFS
Sbjct:   244 KFIDASSTPKVVTFDKNPDNHPYLLKYFQSNAPKAMLFLNFS 285
```

```
>sp|P52589|PDI_WHEAT PROTEIN DISULFIDE ISOMERASE PRECURSOR (PDI) /     983
           DOLICHYL-DIPHOSPHOOLIGOSACCHARIDE-PROTEIN GLYCOTRANSFERASE
           (GLYCOSYLATION SITE-BINDING CHAIN) (GSBP) >gi|508975 (U11496)
           protein disulfide isomerase [Triticum aestivum] >prf||2106410A
           protein disulfide isomerase [Triticum aestivum]
           Length = 515

  Plus Strand HSPs:

 Score = 42 (18.8 bits), Expect = 2.2e-20, Sum P(4) = 2.2e-20
 Identities = 8/8 (100%), Positives = 8/8 (100%), Frame = +1

Query:    37 MAISKVWI 60
             MAISKVWI
Sbjct:     1 MAISKVWI 8

 Score = 183 (83.8 bits), Expect = 2.2e-20, Sum P(4) = 2.2e-20
 Identities = 31/32 (96%), Positives = 32/32 (100%), Frame = +3

Query:   159 LTLHADNFDDAIAXHPFILVEFYAPWCGHCKT 254
             LTLHADNFDDAIAXHPFILVEFYAPWCGHCK+
Sbjct:    42 LTLHADNFDDAIAKHPFILVEFYAPWCGHCKS 73

 Score = 39 (17.4 bits), Expect = 2.2e-20, Sum P(4) = 2.2e-20
 Identities = 7/9 (77%), Positives = 8/9 (88%), Frame = +2

Query:   248 QDLAPEYEK 274
             + LAPEYEK
Sbjct:    72 KSLAPEYEK 80

 Score = 45 (20.6 bits), Expect = 2.2e-20, Sum P(4) = 2.2e-20
 Identities = 9/10 (90%), Positives = 10/10 (100%), Frame = +1

Query:   271 EAAQLLSKHD 300
             +AAQLLSKHD
Sbjct:    80 KAAQLLSKHD 89
```

CSIRO

Gene discovery in the area of wheat quality

- Bioinformatics (storage and analysis of data)

The volume of data now being processed needs an efficient storage and analysis system to compile:

- first pass sequence information
- clone identity and its position within the grided array of library
- record quality of the sequence information

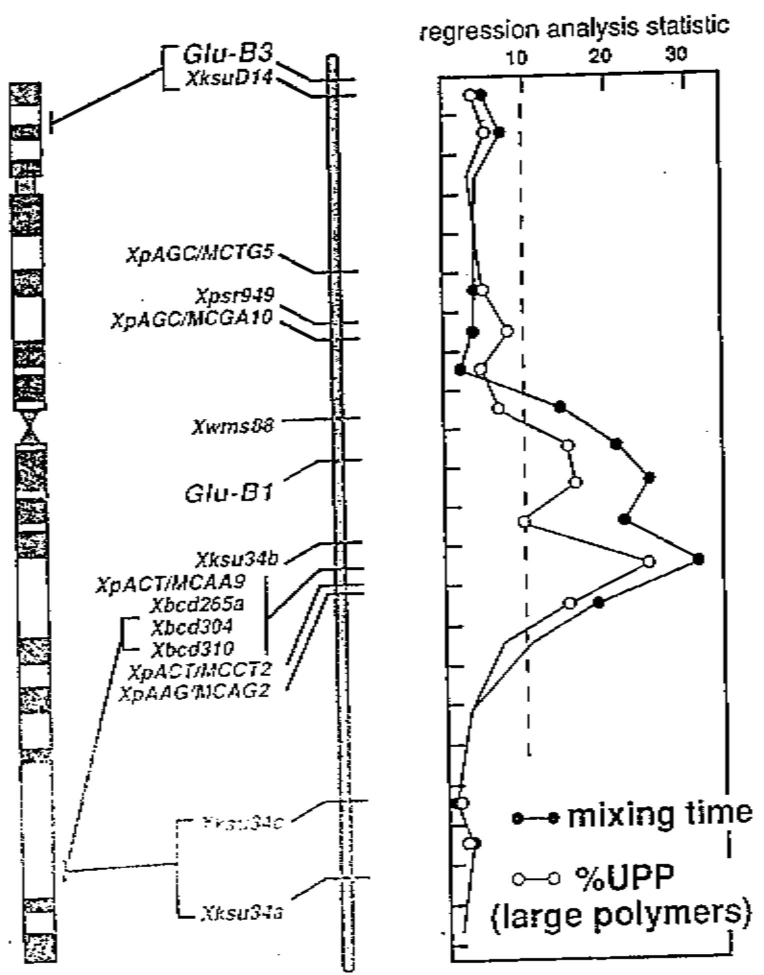
Gene discovery in the area of wheat quality

- Bioinformatics (storage and analysis of data)

(continued)

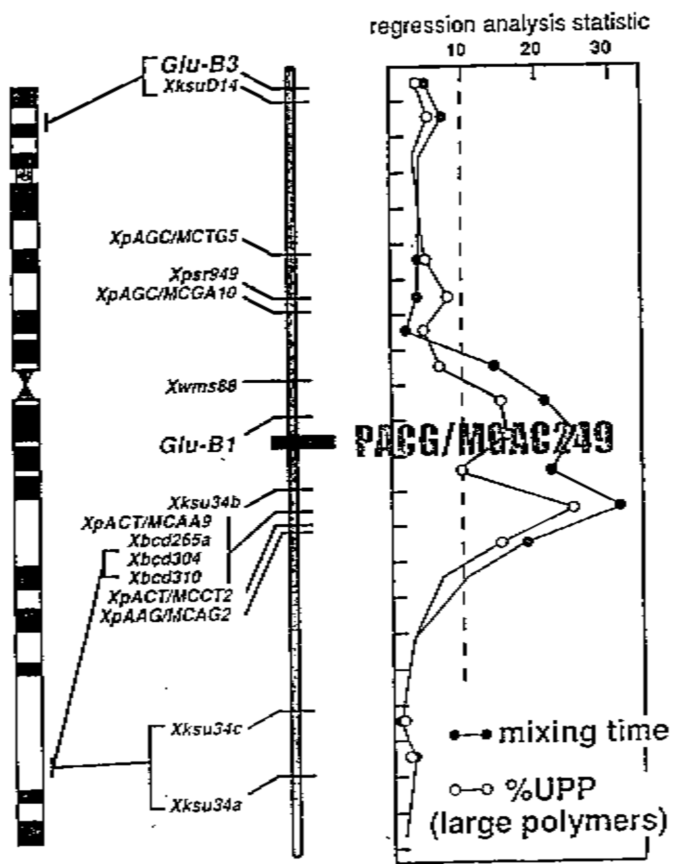
- blast search results for a given sequence, including the P score, best match
- relate sequence back to the full blast results
- full sequence analysis of individual gene family members and the results from sequence alignments
- hybridization data for the time and level of expression of genes

Chromosome 1B of wheat

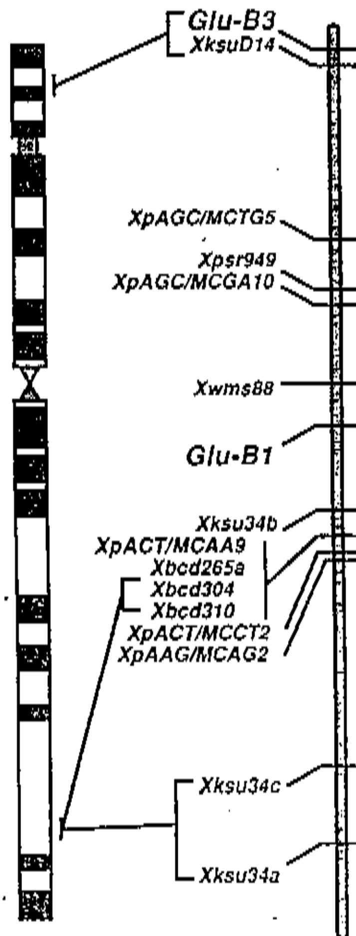


ra98857a-6

Producing a genetic linkage map



The group 1 chromosomes of wheat account for most of the variation in many of the flour processing attributes



ra98857a-3

A selection of storage tissue mutations in cereals and peas*

amylose extender (ae) — Maize mutation giving rise to starch with 60-75% amylose. Locus codes for starch branching enzyme IIb

B10 — Barley mutation exhibiting unorganized callus-like cell populations in addition to the usual starchy endosperm and aleurone.

B7 — Barley mutation which arrests development of grain at syncytial stage.

B9 — Barley mutation lacking an aleurone layer

brittle 1 (bt1) — Maize mutation. Locus codes for adenylate translocator responsible for ... . On chromosome 5L.

brittle 2 (bt2) — Maize mutation. Locus codes for small subunit of ADPG pyrophosphorylase. Equivalent locus in *Chlamydomonas* is *sta 5*. On chromosome 4S.

defective kernel (dek 1) — Maize mutant resulting in failure of aleurone in endosperm development and failure of shoot meristem in embryo development.

dull (du1) — Mutant in maize generating an eponymous dull luster of the kernel. Locus encodes for a soluble starch synthase.

floury2 (fl2) — Maize mutation. Kernels have opaque appearance and are soft and chalky.

M153 — Barley mutant defective in early endosperm development, after cellularisation has occurred.

N17, B13 — Barley mutations that have separate endosperm wings and lack the dorsal prosmatic starchy endosperm cells.

N2 — Barley mutation in which anticlinal cell walls fail to reach the closing stage.

N34 — Barley mutation in which either the left or right endosperm halves fail to develop normally.

opaque 2 (o2) — Mutant in maize resulting in a reduction in α-zein content. The locus codes for an endosperm-specific member of the basic leucine zipper (bZIP) family of transcription factors

rugosus3 (rug3) — Pea mutant giving a wrinkled seed phenotype. Locus encodes for plastidial phosphoglucomutase.

rugosus5 (rug5) — Mutant in pea giving a wrinkled seed phenotype. Locus encodes for a soluble starch synthase and equivalent locus in *Chlamydomonas* is *sta3*.

shrunken 1 (sh1) — Maize mutant giving a wrinkled seed phenotype. Locus encodes for sucrose synthase 1 (SS1) - equivalent locus in peas is *rugosus4 (rug4)*. On chromosome 9S of maize)

shrunken2 (sh2) — Maize mutation giving a shrunken kernel. Locus codes for large subunit of ADPG pyrophosphorylase. Equivalent locus in *Chlamydomonas* is *sta 1*. On chromosome 3L.

soft starch (h1) — Maize mutation with endosperm soft and opaque. On chromosome 3.

sugary 1 (su1) — Maize mutation with increased level of sugar in kernel. Locus encodes for starch debranching enzyme (isoamylase). On chromosome 4S

sugary 2 (su2) — Maize mutation associated with glassy endosperm, slightly shrunken phenotype of kernel and altered granule morphology. On chromosome 5L.

sucrose synthase1 (sus 1) — Maize mutant - locus codes for sucrose synthase 2 (SS2).

viviparous 1 (vp1) — Maize mutant where kernels germinate before ear matures. Due to insensitivity to abscisic acid. Locus encodes a transcriptional activator which interacts synergistically with ABA-regulated transcription factors.

waxy (wx) — Mutants in maize that give the kernel a waxy appearance. They are devoid of amylose starch due to absence of the granule bound starch synthase (GBSS). Similar mutants are found in wheat, barley, *Chlamydomonas (sta2)*, pea (*lam*) and potato. On chromosome 9S.

* data based on Doan et al (1996), Neuffer et al (1997), Wang and Cuming, (1995), Wang et al (1998)

Heller, RA., Schena, M., Chai, A., Shalon, D., Bedilion, T., Gilmore, J., Woolley, DE., Davis, RW. 1997. Discovery and analysis of inflammatory disease-related genes using cDNA microarrays. Proc. Natl. Acad. Sci. USA 94, 2150 - 2155.

Huber AG., Grabe DF. 1987. Endosperm morphogenesis in wheat transfer of nutrients from the antipodals to the lower endosperm. Crop Science. 27, 1248-1252.

Huber AG., Grabe DF. 1987. Endosperm in wheat: termination of nuclear division. Crop Science. 27, 1252-1256

Huttly, AK., Martienssen, RA., Baulcombe, DC. 1988. Sequence heterogeneity and differential expression of the α-Amy2 gene family in wheat. Mol. Gen. Genet. 214, 232 - 240.

Johnston, M. 1998. Gene chips: Array of hope for understanding gene regulation. Current Biol. 8, R171 - R174.

Joshi,C.P. and Nguyen,H.T. 1994. Differential display mediated rapid cloning and sequencing of the 3' region of several members of a large multigene family. GenBank WHTHSP169D.

Ko, MSH., 1990. An equalised cDNA library by the reassociation of short double-stranded cDNAs. Nucl. Acids Res. 18, 5705 - 5711.

Kohchi, T., Fujishige, K., Ohyama, K. 1995. Construction of an equalised cDNA library from *Arabidposis thaliana*. The Plant J. 8, 771 - 776.

Kreis, M., Shewry, PR., Forde, BG., Forde, J., Miflin, BJ. 1985. Structure and evolution of seed storage proteins and their genes with particular reference to those of wheat, barley and rye. In Oxford Surveys of Plant Molecular and Cell Biology. B. Miflin ed. Oxford University press, Oxford UK. Pp 253 - 317.

Lagudah, ES., Moullet, O., Appels, R. 1997. Map-based cloning of a gene encoding a nucleotide-binding domain and a leucine-rich region at the *Cre3* nematode resistance locus of wheat. Genome 40, 659 - 665.

Lazo, GR., Larka, LA., Hsia, CC., McCue, KF., Sorrells, M., Matthews, DE., Au, M., Federspiel, NA., Anderson, OD. 1998. Assigning putative gene functions to mapped probe loci in the GrainGenes genome database and sequencing of wheat endosperm cDNAs. Plant and Animal Genome VI Conference, January 18-22, 1998, San Diego, CA, USA.

Lersten, NR. 1987. Morphology and anatomy of the wheat plant. In Wheat and wheat improvement. 2nd edition, ed EG. Heyne. American Society of Agronomy monograph. p33- 78.

Levanony, H., Rubin, R., Altschuler, Y., Galili, G. 1992. Evidence for a novel route of wheat storage proteins to vacuoles. J. Cell Biol. 119, 1117 - 1128.

Litts, JC., Colwell, GW., Chakerian, RL., Quatrano, RS. 1991. Genomic nucleotide sequence of a *Triticum aestivum* 7S globulin (Gbl 1). GenBank WHTGBL1A.

Liu, J., Hara, C., Umeda, M., Zhao, Y., Okita, TW., Uchimiya, H. 1995. Analysis of randomly isolated cDNA's from developing endosperm of rice (*Oryza sativa* L): evaluation of expressed sequence tags and expression levels of mRNAs. Plant Mol. Biol. 29, 685 - 689.

Long, JA., Moan, E.I., Meford, JI., Barton, MK. 1996. A member of the *KNOTTED* class of homeodomain proteins encoded by the *STM* gene of *Arabidopsis*. Nature 379, 66-69.

López-Castañeda, C., Richards, R.A., Farquhar, G.D., and Williamson, R.E. 1996. Seed and seedling characteristics contributing to variation in early vigor among temperate cereals. Crop Science 36, 1257-1266.

Marana,C., Garcia-Olmedo,F. and Carbonero,P. 1988. Linked sucrose synthase genes in group-7 chromosomes in hexaploid wheat (Triticum aestivum L.). Gene 63, 253-260

Mares DJ., Norstog K., Stone B A. 1975. Early Stages in the development of wheat endosperm. 1. The change from free nuclear to cellular endosperm . Australian Journal of Botany . 23, 311-326

Mikami,K., Sakamoto,A. and Iwabuchi,M. 1994. The HBP-1 family of wheat basic/leucine zipper proteins interacts with overlapping cis-acting hexamer motifs of plant histone genes. J. Biol. Chem. 269 (13), 9974-9985.

Mrva, K., Mares, DJ. 1996. Inheritance of late maturity α-amylase in wheat. Euphytica 88, 61 - 67.

Muller, M., Knudsen, S. 1993. The nitrogen response of a barley C-hordein promoter is controlled by poistive and negative regulation of the GCN4 and endosperm box. The Plant J. 4, 343 - 355.

Murai, N., Li, Z., Kawagoe, Y., Hayashimoto, A. 1991. Transposition of the maize activator element in transgenic rice plants. Nucl. Acids Res. 19, 617 - 622.

Nair, R., Baga, M., Scoles, GJ., Kartha, K., Chibbar, R. 1997. Isolation, characterisation and expression analysis of a starch branching enzyme II cDNA from wheat. Plant Sci. 122, 153 - 163.

Napier,J.A., Shewry,P.R., Grimwade,B. and Freedman,R.B. 1994, Genbank X79306

Neuffer, MG., Coe, EH., Wessler, SR. 1997. Mutants of Maize. Cold Spring Harbor Laboratory Press, 486 pp.

Niemietz C., Jenner CF. 1993. Mechanisms of sugar into endosperm and aleurone protoplasts isolated from developing wheat germs. Australian Journal of Plant Physiology. 20, 371-378.

Olive,M.R., Ellis,R.J. and Schuch,W.W. Isolation and nucloetide sequences of cDNA clones encoding ADP-glucose pyrophosphorylase polypeptides from wheat leaf and endosperm. Mol. Biol. 12, 525-538 (1989)

Ouellet,F., Vazquez-Tello,A. and Sarhan,F. 1998. The wheat wcs120 promoter is cold-inducible in both monocotyledonous and dicotyledonous species. FEBS Lett. 423, 324-328.

Patanjali, SR., Parimoo, S., Weissman, SM. 1991. Construction of a uniform-abundance (normalized) cDNA library. Proc. Natl. Acad. Sci. USA 88, 1943 - 1947.

© 1999 Oxford University Press

Nucleic Acids Research, 1999, Vol. 27, No. 2 601–607

'LABNOTE', a laboratory notebook system designed for academic genomics groups

Marie-Christine Imbert[1], Van Khanh Nguyen, Samuel Granjeaud, Catherine Nguyen and Bertrand R. Jordan*

TAGC group, ICIM, Centre d'Immunologie INSERM/CNRS de Marseille-Luminy, Case 906, 13288 Marseille Cedex 9, France and [1]Centre de Thérapie Génique, IPC, 232 Boulevard de Ste Marguerite, 13273 Marseille Cedex 9, France

Received August 31, 1998; Revised and Accepted November 20, 1998

ABSTRACT

We have developed a relational laboratory database system, adapted to the daily book-keeping needs of laboratories that must keep track of information acquired on hundreds or thousands of clones in an effective and user-friendly fashion. Data, whether final or related to experiments in progress, can be accessed in many different ways, e.g. by clone name, by gene, by experiment or through DNA sequence. Updating, import and export of results is made easier by specially developed tools. This system, in network version, serves several groups in our Institute and (over the Internet) elsewhere, and is instrumental in collaborative studies based on expression profiling. It can be used in many similar situations involving progressive accumulation of information on sets of clones or related objects.

INTRODUCTION

Many public databases have been established to store and make available all kinds of genomic data, from maps to sequences through catalogues of mutants and protein motifs (1). Recent efforts have been aimed, in particular, at making gene expression data publicly available and at the same time providing users with data analysis tools (2). In addition, large laboratories, such as Genome Centers involved in intensive genome mapping or sequencing, have set up their own database systems to store their results and prepare the data for distribution to the scientific community. Being geared to a particular operation, such systems are rarely made available or even published. One of them, the 'Genome Notebook', developed primarily to handle results from the human chromosome 11 mapping project, has however been described in some detail (3).

Increasingly, 'conventional' laboratories (i.e. groups of relatively small size operating in an academic environment) are interfacing with the genome project and making use of its results (e.g. sequence or mapping data), but also of resources such as the IMAGE cDNA clone set (4), and of semi-automated procedures that boost throughput by one or two orders of magnitude. This trend results in a large increase in the number of objects and in the amount of information, requiring efficient archiving and easy retrieval of experiments in progress, intermediate results and 'final' data. The traditional approach, i.e. manual notebooks supplemented by a number of computer files in spreadsheet software, can no longer cope with this data flow, and a proper laboratory database becomes necessary.

A number of projects developed in our Institute are centred around genes expressed in the mouse thymus. We use organised cDNA libraries, measure expression levels by hybridisation of DNA arrays with complex probes, and obtain additional information (tag sequence, genome mapping, etc.) for sets of clones selected according to their expression pattern (5–8). Thus the information we wish to store in a laboratory database is largely organised around a list of clones (expression data, sequences, results from Southerns and northerns, etc.) but also includes the description of libraries, the make-up of specific arrays, as well as protocols or publication references. Other ongoing research programmes use sets of a few hundred IMAGE cDNA clones as reagents for expression profiling in various situations; again, good book-keeping is essential to keep track of clone choice, procurement, verification and of expression data. Ready-made membranes provided by several suppliers (http://www.clontech.com/clontech/Catalog/Hybridization/Atlas.html , http://www.genomesystems.com/GDA/ and http://www.resgen.com/) as well as by resource centres are also used in some projects and generate a need for data archiving.

To be really useful in the context of a biological laboratory, a notebook system must be extremely user-friendly: it should be used daily by each member of the group, and the interface must be designed with this in mind. It should run well on affordable machines that the prospective users are familiar with, i.e. in most cases on PC or Macintosh microcomputers. The system must be extremely flexible and allow additions and changes to be made without loss of previously entered data, to accommodate easily new experimental approaches or new ways of analysing existing information; data security and access privileges should also be well organised.

We have used the 4th Dimension (ACI) relational database management system (http://www.aci-4D.com/) to develop a laboratory database, LABNOTE, aiming to fulfil this need. 4th Dimension (4D) has been used previously for biological databases (9,10), for a large number of (unpublished) medical databases,

*To whom correspondence should be addressed. Tel: +33 4 91 26 94 96; Fax: +33 4 91 26 94 30; Email: jordan@ciml.univ-mrs.fr

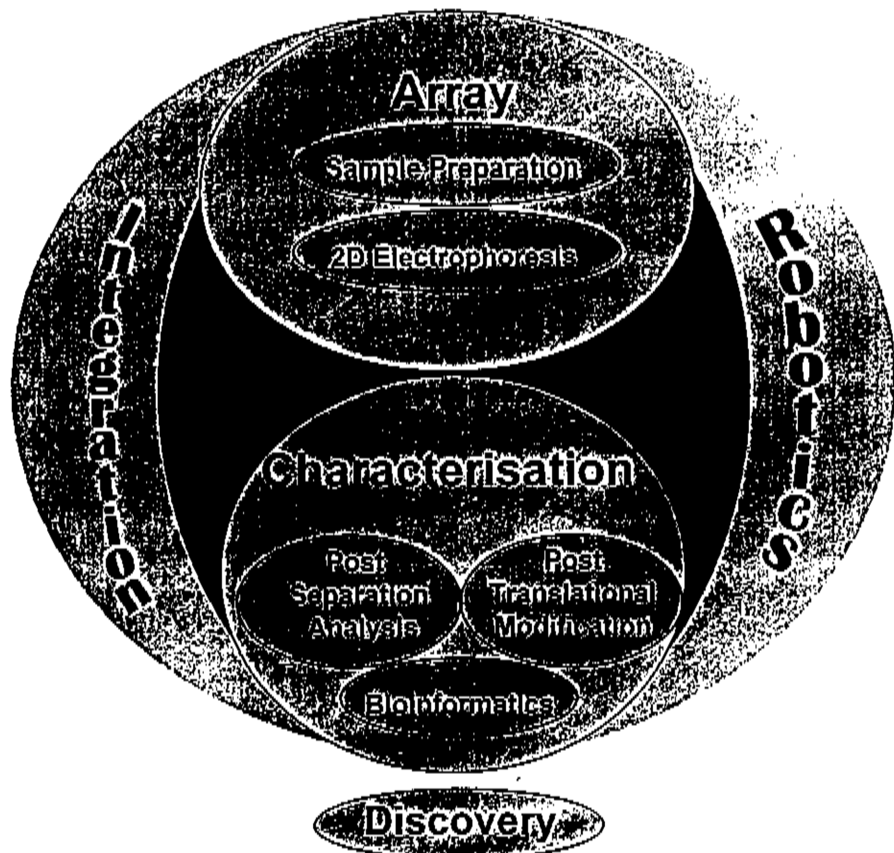
Strategic Alliances

PIONEER HI-BRED INTERNATIONAL, INC.

What do these places have in common ?

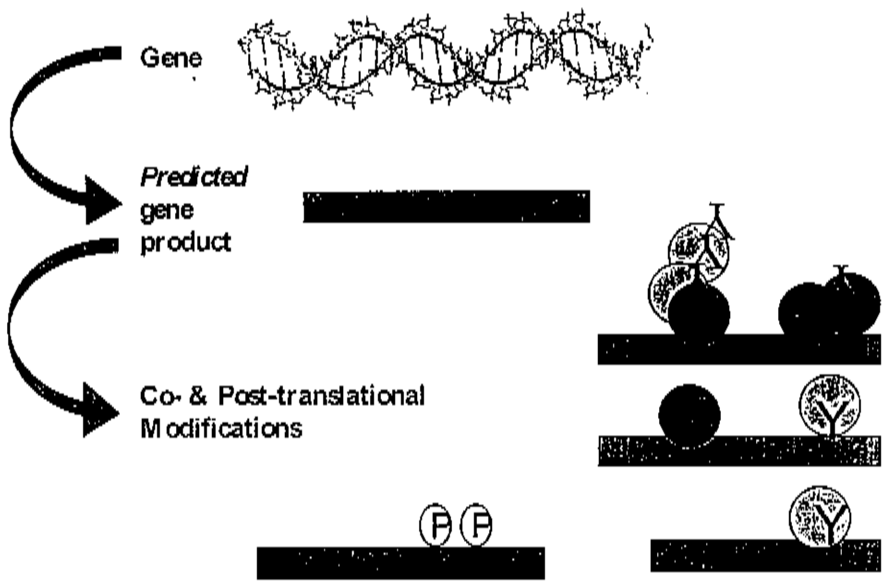
Uppsala, Sweden

Hercules, California

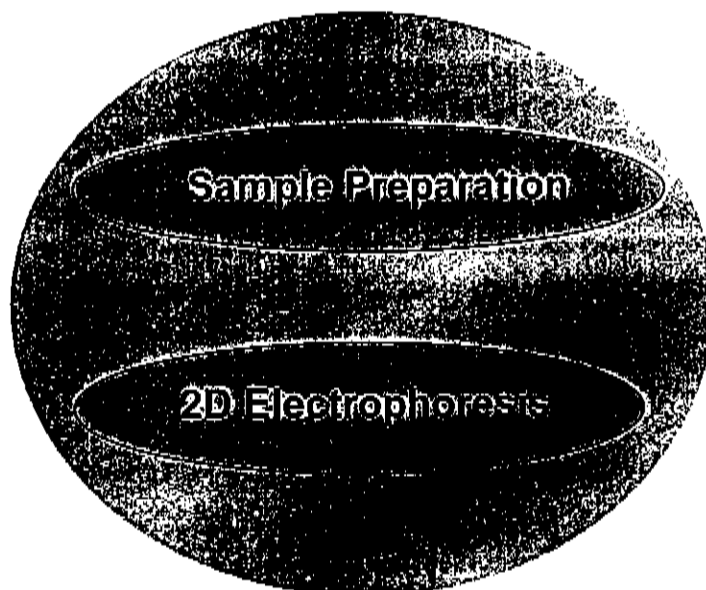


Why Proteins ?

One gene ➡ Many proteins



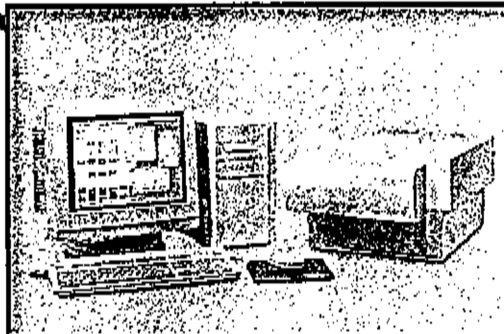
Array



Sample Preparation

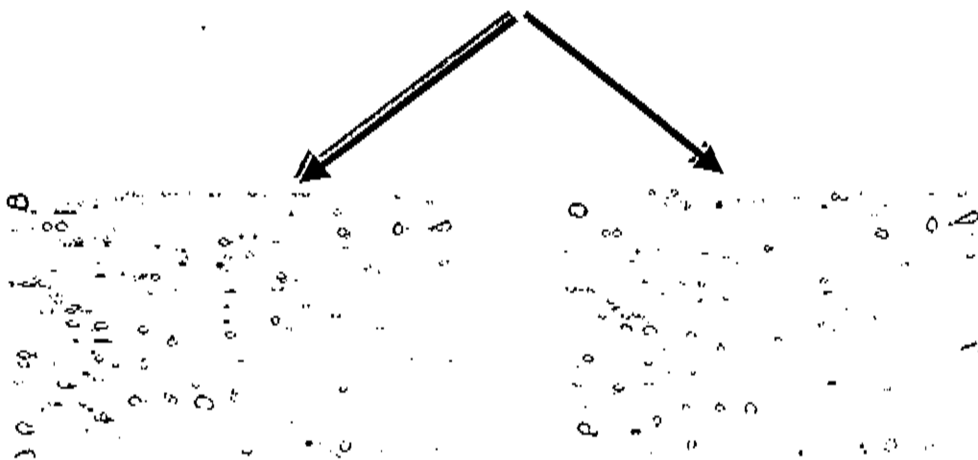
- Most critical step
- Varies according to sample
  - Precipitation to remove salts and/or contaminants
  - Soluble samples
- Prefractionation

Imaging and Melanie



Differential Protein Display

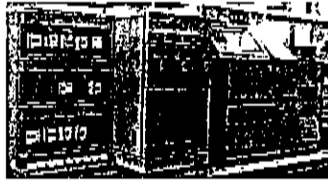
Silver stained gels analysed using
Melanie II



Control vs. test

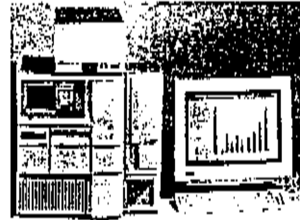
Post Separation Analysis

MALDI-TOF mass spectrometry



•N & C terminal sequencing

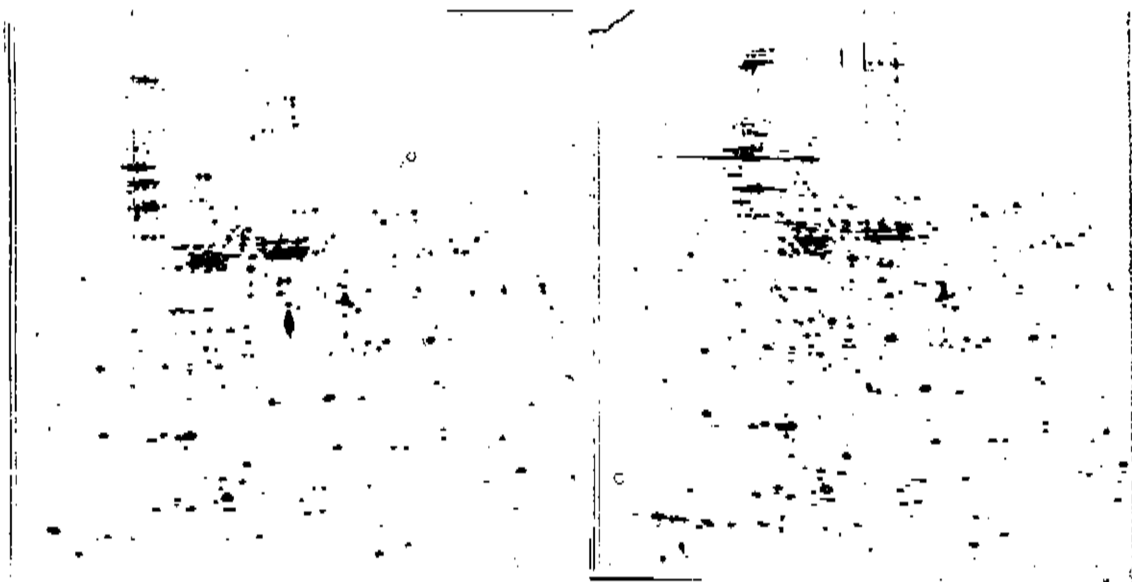
•Amino Acid Composition



•Bioinformatics

expasy.proteome.org.au

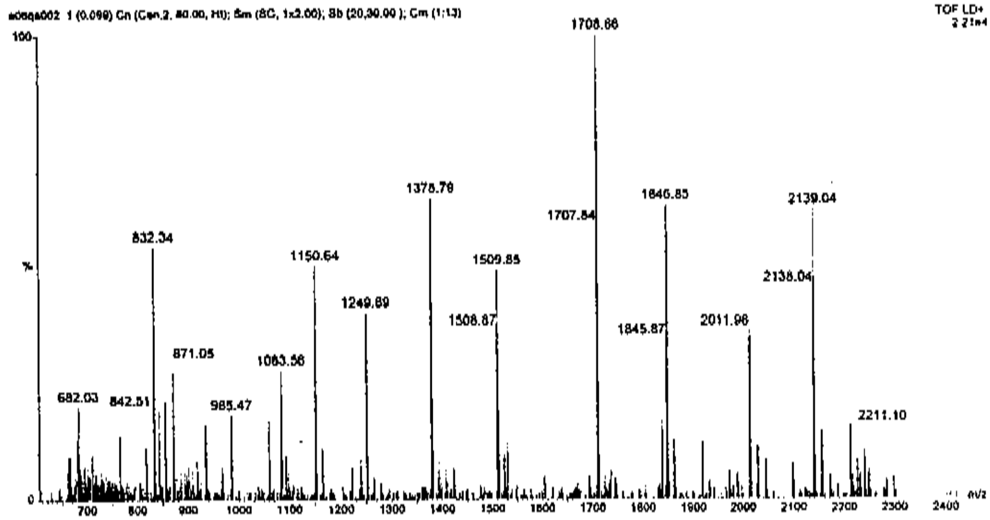
Increasing Loads of *E. coli.*



One milligram Three milligrams

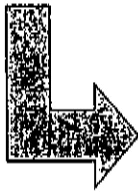
Colloidal Coomassie Blue stained

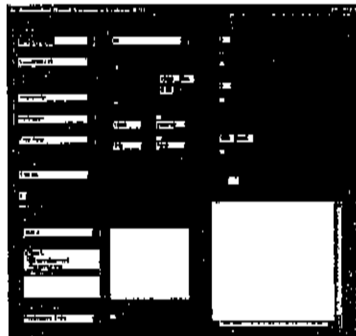
MALDI spectrum of E. *coli* protein spot



Peak table to RAID

765.05	934.53	1377.78	2011.98
817.45	985.47	1508.87	2028.02
832.34	1060.06	1529.85	2138.04
842.51	1083.56	1707.84	2154.02
855.05	1150.64	1845.87	2211.10
871.05	1249.69	1860.86	





P02339

Summary Statistics

Date Started	12/6/98
Number of gels	3
No. of spots analysed with MALDI	1085
No. of spots confidently identified	565
No. of spots putatively identified	55
No. of unique proteins	195
No. of spots identified on the 3mg master	324
No. of spots on 1mg master not on 3mg	17
No. of spots on 6mg master not on 3mg	21

N and C Terminal Sequence

ASRGVNKVIL VGNLGQDPEV RYMPNGGAVA NITLATSESW RDKATGEMKE QTEWHRVV
GKLAEVASEY LRKGSQVYIE GQLRTRKWTD QSGQDRYTTE VVVNVGGTMQ MLGGRQG
PAGGNIGGGQ PQGGWGQPQQ PQGGNQFSGG AQSRPQQSAP AAPSNEPPMD FDDDIPF

Post Translational Modifications

FindMod and MALDI-TOF MS PSD

Glycoprotein detection from blots
Glycosite sequencing

Phospho AAA
Monosaccharide profiling

Bioinformatics

FindMod

STHM_HUMAN (P16949)
STATHMIN (PHOSPHOPROTEIN P19) (PP19) (ONCOPROTEIN 18) (OP18) (LEUKEMIA-ASSOCIATED PHOSPHOPROTEIN P18) (PP17) (PROSOLIN) (METABLASTIN) (PR22 PROTEIN).
HOMO SAPIENS (HUMAN).

Theoretical pI/Mw: 5.77 / 17171.32

Calculate
the theoretical masses of peptides generated by the chemical or enzymatic cleavage of this protein using PeptideMass.

Matching peptides:

User mass	DB mass	#MC	peptide	position	known modifications
1388.75000	1388.72731	1	ASSDIQVKELEK	1-12	(ACET: 1)
1388.75000	1388.75376	0	ASGQAFELILSPR	14-26	
1541.80000	1541.8215	1	SKESVPEFPLSPPK	27-40	
1326.73000	1326.69451	0	ESVPEFPLSPPK	29-40	
945.48800	945.50051	1	KLEAAEER	52-59	
1165.59000	1165.54893	0	AIEENNNFSK	85-94	

Potentially modified peptides, detected by mass difference and conforming to rules (considering only peptide masses that have not matched above):

User mass	DB mass	mass diff.	mod. diff.	potential mod.	#MC	peptide	position	known modifications
1410.78000	1330.75818	80.02182	79.96630	PHOS	2	KKDLSLEEIQK	41-51	
1410.78000	1330.75818	80.02182	79.96630	PHOS	2	KDLSLEEIQKK	42-52	

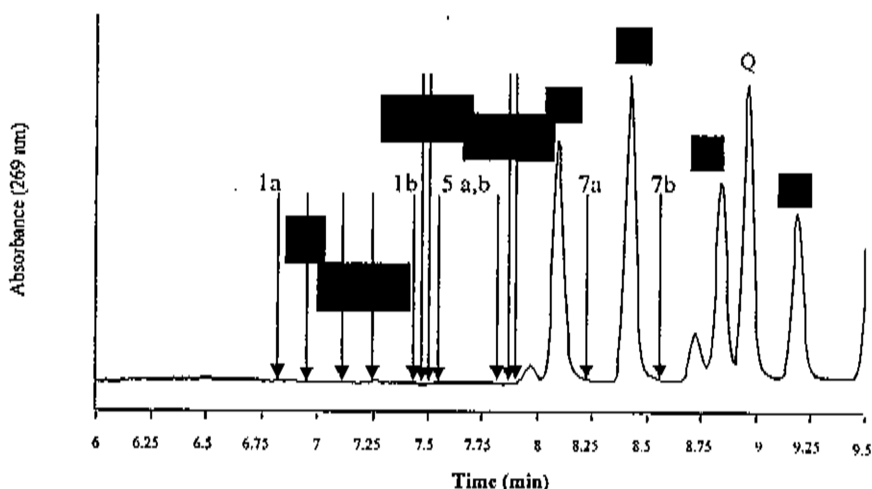
Phosphorylation

Kingdom	Residues	Position	Rule
Eukaryotes, Viruses	S,T,Y,H,D	anywhere	-
Prokaryotes	S,T,H,C,D	anywhere	-

Mass difference:

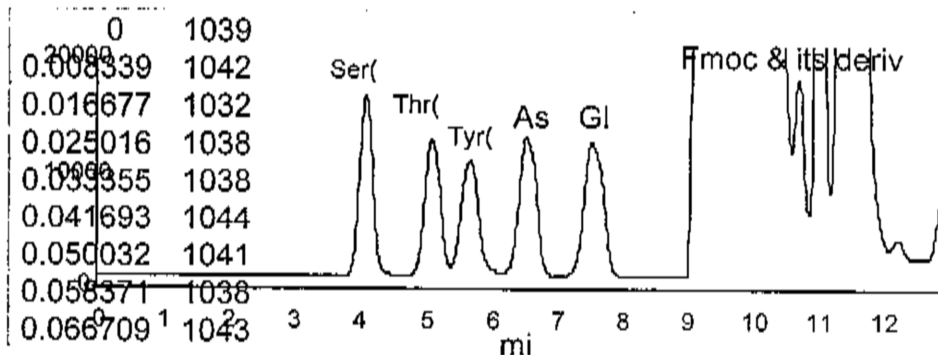
average	79.9799
monoisotopic	79.9663

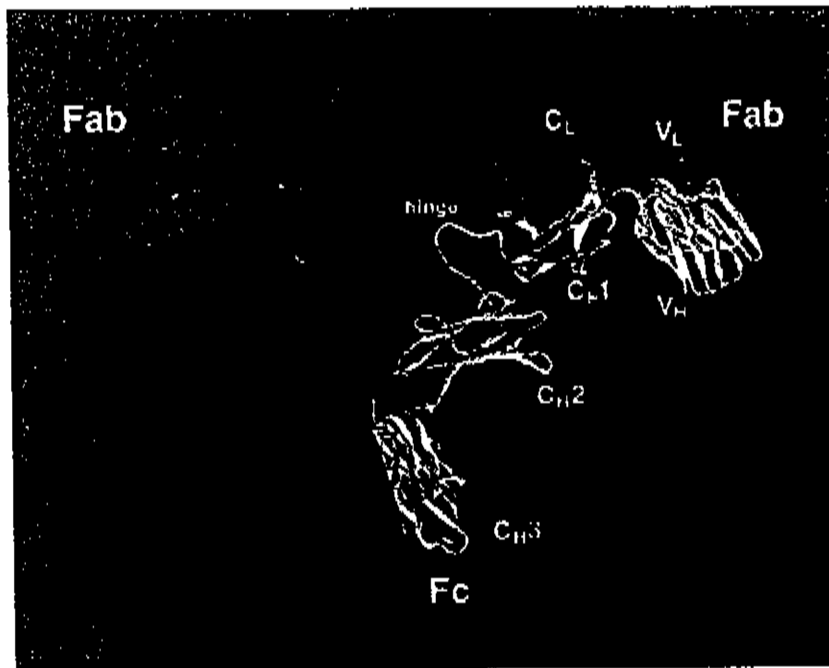
GlycoSite™ PTH-glycoamino acid Profiles

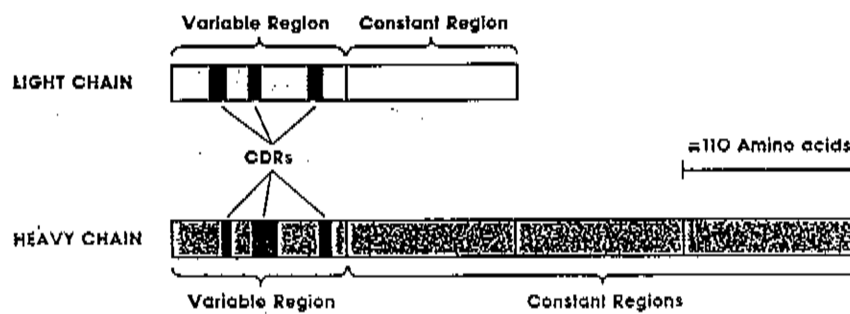


1a,b	Thr-HexHex	2	Ser-HexHex
5a,b	Thr-HexNAcHex	3a,b	Ser-Hex
7a,b	Thr-HexNAc	4a,b	Ser-HexNAcHex
		6a,b	Ser-HexNAc

Separation of three standard phosphoamino acids







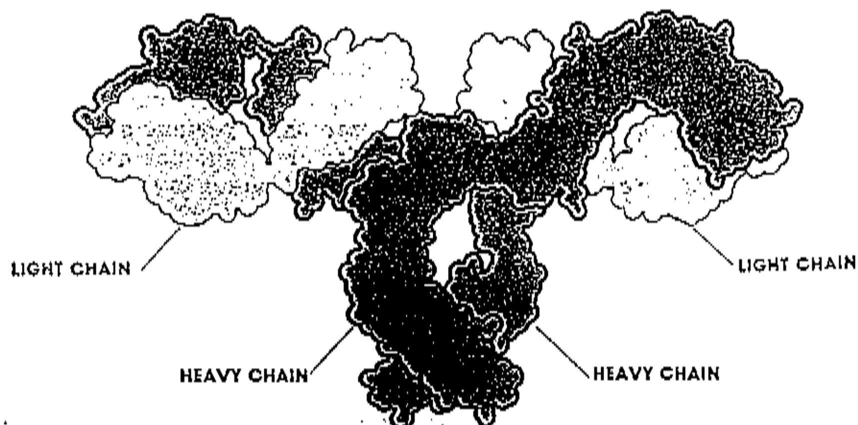


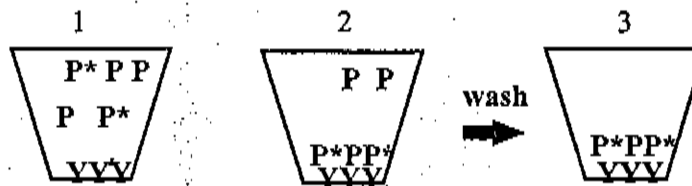
FIGURE 2.5
Light- and heavy-chain structure. (Adapted from Silverton et al. 1977.)

Competitive enzyme-immunoassay

CSIRO PLANT INDUSTRY

CRC

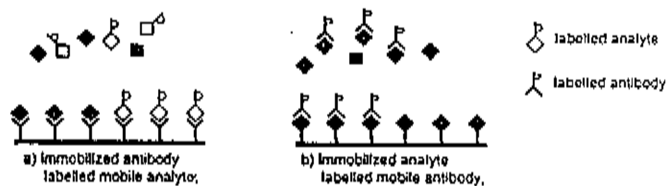
1. Sample (containing pesticide, P) and enzyme-pesticide conjugate (P*) added to antibody-coated microwell. Incubated 5-60 min.
2. Competition between P and P* for limited number of antibody binding sites. Unbound pesticide and enzyme conjugate removed by washing step.
3. Colour developer (enzyme substrate/chromogen) added.



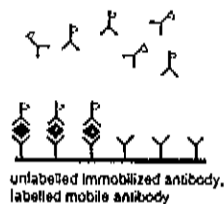
Increasing concentrations of pesticide produce decreasing colour

1) occupancy of antibody binding sites by the analyte

2) competitive assays (limiting antibody concentration)



3) non-competitive assays (antibody excess)



unlabelled immobilized antibody, labelled mobile antibody

CSIRO PLANT INDUSTRY

WHEAT CRC

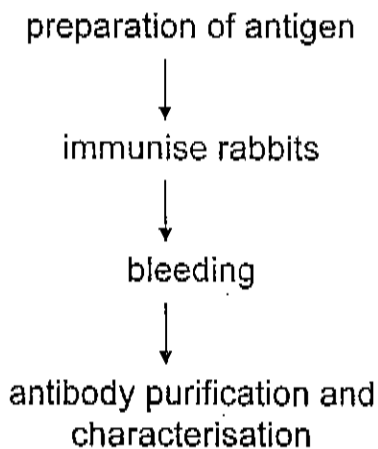
Advantages of immunoassays

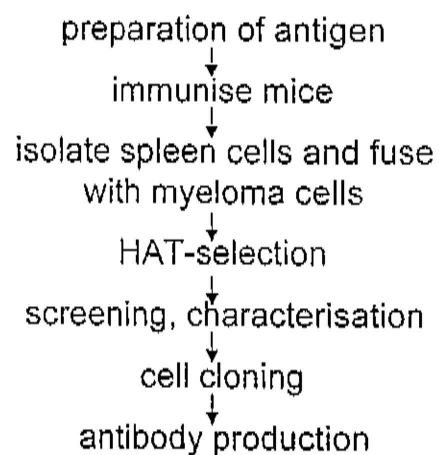
- Speed
- High sample throughput (96-well microtiter plate format)
- No need for special skills
- No need for expensive equipment
- Can be developed for almost any component
- low amount of samples

CSIRO PLANT INDUSTRY

WHEAT CRC

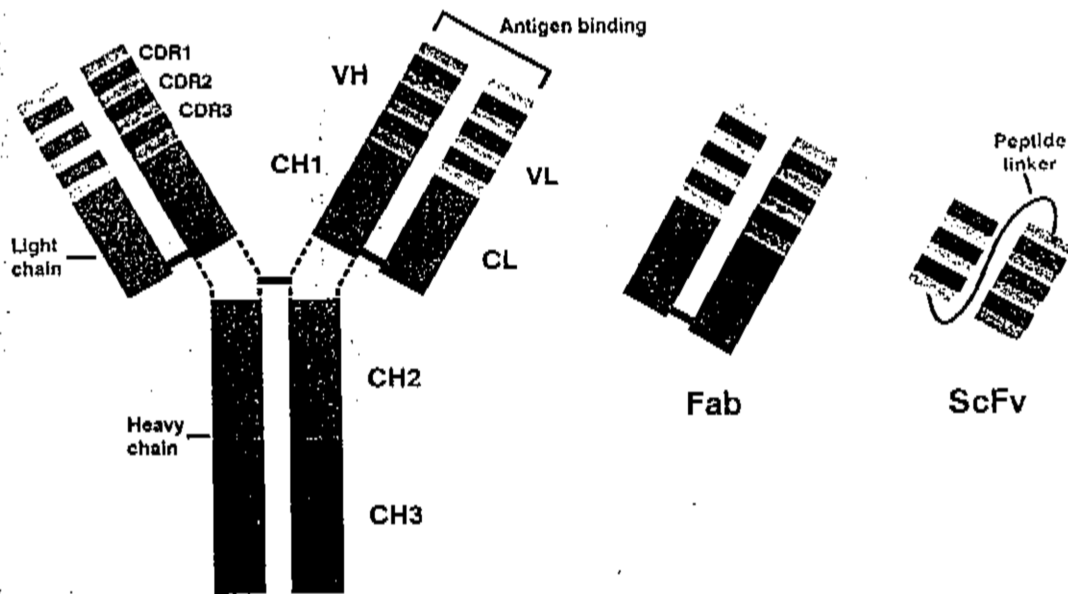
Antibody production





PLANT INDUSTRY

ANTIBODY STRUCTURE



CL97753A

CSIRO PLANT INDUSTRY

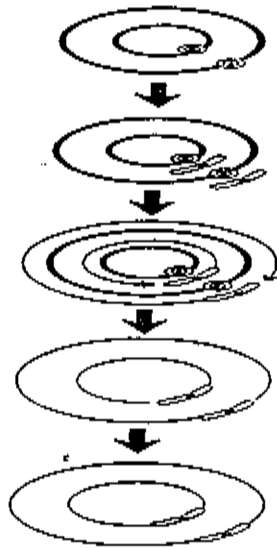
Steps from hybridoma/spleen cells to phage libraries

- isolation of mRNA
- reverse transcription into cDNA
- separate PCR amplification of variable heavy- and light chain regions
- assembly of H- and L-chain with linker (ScFv)
- Sfi I / Not I - digest
- cloning into phagemid pCANTAB 5E
- transformation of E. coli (TG1)
- screening → enrichment by biopanning
  → colony lift assay

CSIRO PLANT INDUSTRY

Overview of site-directed mutagenesis

WHEAT CRC



Gene in plasmid with target site for mutation

Denature plasmid and anneal mutagenic primers

Generate non-methylated mutant copy

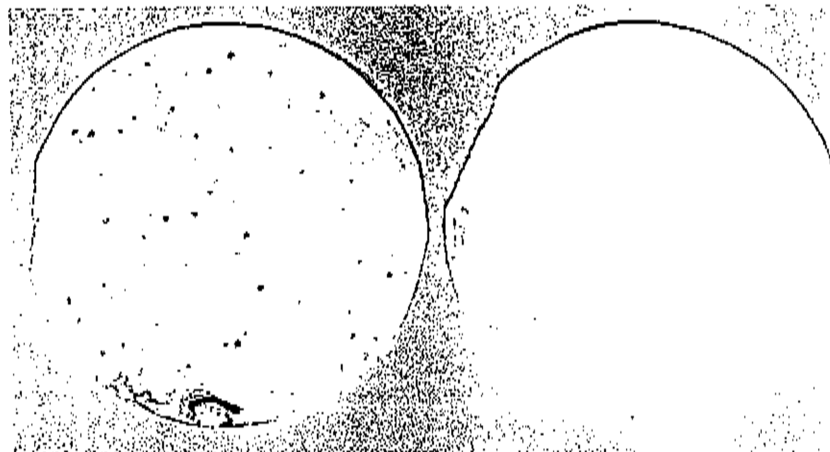
Digest the methylated, nonmutated parental DNA with Dpn 1

Transform mutated DNA into supercompetent cells

CSIRO PLANT INDUSTRY

Colony lift assay with 110622 library (after 1 round panning)

WHEAT CRC

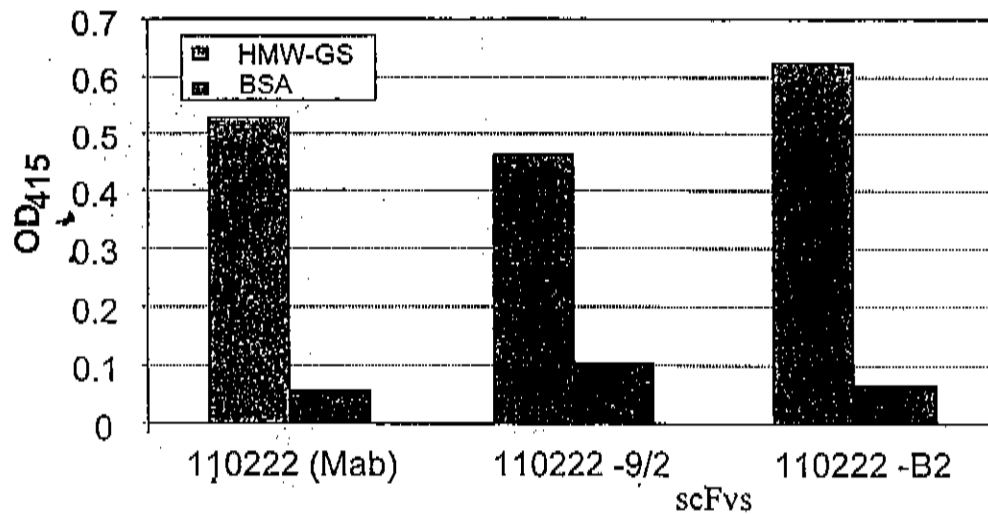


membrane coated with: HMW-GS NFDM

CSIRO PLANT INDUSTRY

ELISA with scFvs and mabs

WHEAT CRC



CSIRO PLANT INDUSTRY

Proposed initiative: High-throughput DNA-based diagnostics

WHEAT CRC

Aim: Delivery of molecular markers as high-throughput ELISA-like diagnostic tests

Outcome: Accurate and rapid identification of key quality and productivity traits for breeding programs

Strategies: Compare PCR-based amplification methods with DNA - capture detection (Direct hybridisation) techniques

Targets : Genes for specific quality traits
- late maturity alpha-amylase, dormancy genes
- starch synthetic enzymes
- polyphenol oxidases

Key storage protein (gliadin and glutenin) alleles
Hardness/water absorption markers
Disease resistance markers (*Sr 2*, *Sr 26*, *Cre-3*)

CSIRO PLANT INDUSTRY

Strategies to alter specificity/affinity of recombinant antibodies

CRC

- mutation: site directed --> PCR-based
  random --> mutator strains
- create libraries from naïve or immunised spleen cells
- increase diversity by chain shuffling (combined with mutation)

highly diverse libraries need powerful screening systems

CSIRO PLANT INDUSTRY

Phage Display Technology

CRC

- antibody displayed on the surface of phage is highly accessible for binding assays without disrupting the phage particle
- selection by panning is efficient and rapid
  - enrich specific binders from a background irrelevant clones
  - discriminate between high and low affinity binders
  - large numbers of clones may be characterised
  - phage recovery and expansion enable isolation of rare clones
- physical linkage between phenotype and genotype

CSIRO PLANT INDUSTRY

Shortcomings of monoclonal antibodies

WHEAT CRC

- cell fusion during hybridoma production low in efficiency
- some specificities hard to achieve (homologous subunits)
- specificity and affinity of obtained antibodies can not be manipulated

CSIRO PLANT INDUSTRY

Potential of recombinant antibodies

WHEAT CRC

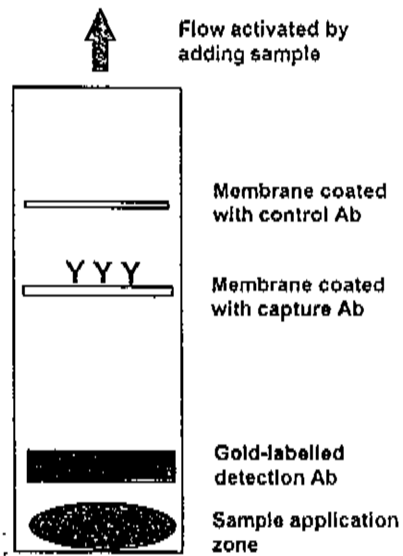
- PCR enabled facile and rapid isolation of antibody sequences (cDNA)
- Phage display and panning is a powerful display and selection method for specific antibodies.
- Potential to manipulate specificity and affinity of of existing antibodies by mutation or chain recombination.
- extend the immune repertoire by creating diverse combinatorial libraries.
- engineering fusion proteins possessing antigen binding sites and labels (enzymes).
- cheap large-scale production in bacteria.
- transfer to different expression systems (insect, mammalian, plant cells.

PLANT INDUSTRY

Immunochromatography field tests

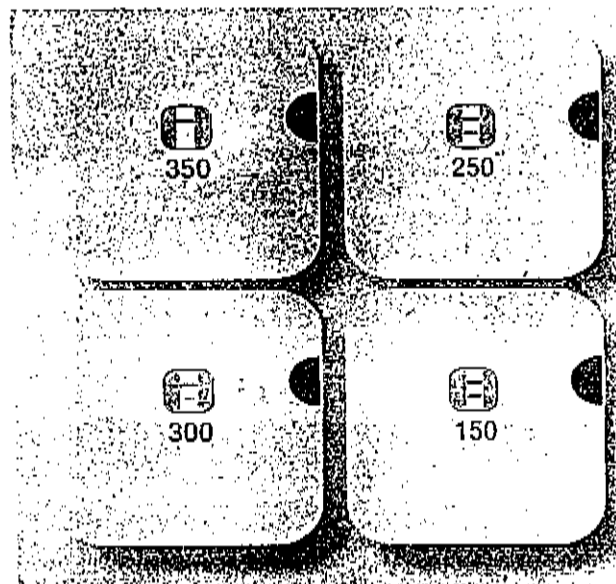
1. Extract grain by shaking in salt solution
2. Apply sample to lower zone; starts upwards movement of solution front
2. Complexes of target antigen in sample and specific gold-labelled antibody form and move upwards
3. Complexes captured by capture antibody. Pink-coloured band forms
4. Control (anti-species) antibody binds excess unlabelled gold-labelled Ab

Positive test : two bands
Negative test : one band (control only)
Test time: 3-7 minutes



CSIRO
PLANT INDUSTRY

Detection of Pre-harvest sprouting using a 5-minute immunochromatography test



Immunodiagnostics in the wheat production and processing chain

Breeder -- Farmer -- Grain storage -- Milling -- Processing -- Consumer

Rain damage test

Chromosome translocations

Gluten in products

Pesticide residues

Dough strength

Mycotoxins

LMAA

HMWGS + LMWGS alleles

Field detection of pesticides and mycotoxins in stored grain

- Enzyme / immunoassay field tests have been developed for most insecticides and mycotoxins present in stored commodities
  methods applied to wheat, barley, milling fractions, products
  tests still required for *Alternaria* toxins

- Assays can detect either individual compounds or groups
  depending on the chemistry used for hapten synthesis

- Tests can either be used as screening tests or for quantification

- Must validate immunoassay against instrumental method using:
  residues extracted with the same solvent/ method
  both incurred and spiked residues for each matrix

Immunodiagnostics for the wheat industry
Impact of gene technology

Thomas M. Giersch, Amanda S. Hill, Barry McCarthy and John H. Skerritt

- Principles of immunodiagnostics
- Immunodiagnostics in the wheat production/processing chain
- Imunoassay formats
- Advantages of immunodiagnostics
- Potential of engineered antibodies
- From antibody-producing cells to recombinant antibodies
- Mutation and screening

Principles of Immunodiagnostics

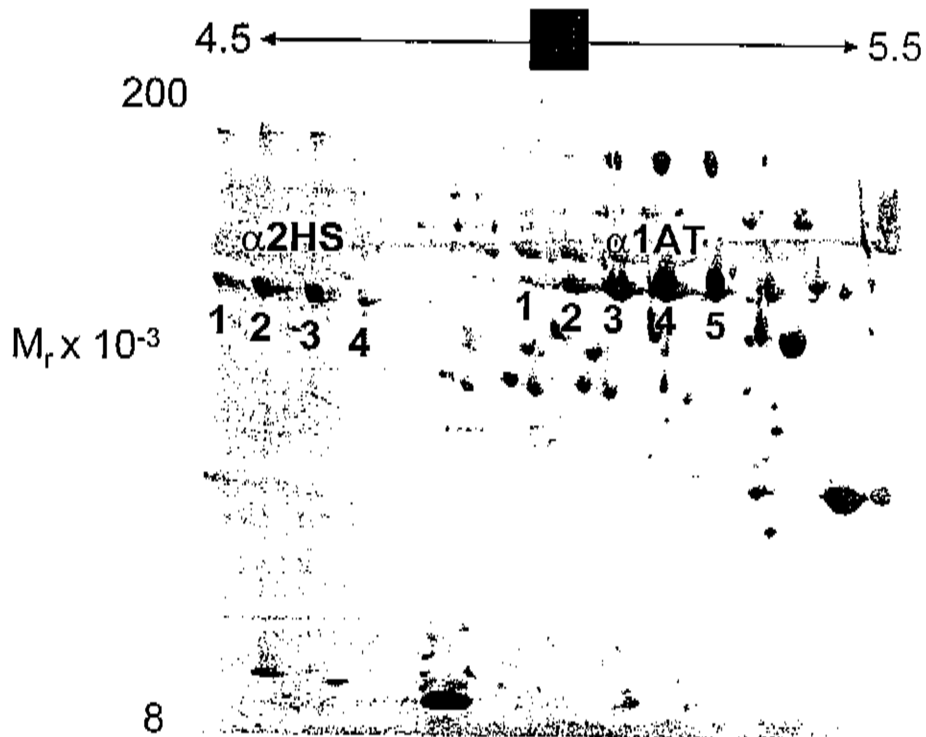
Use of discriminating power of antibodies to detect or quantify substances

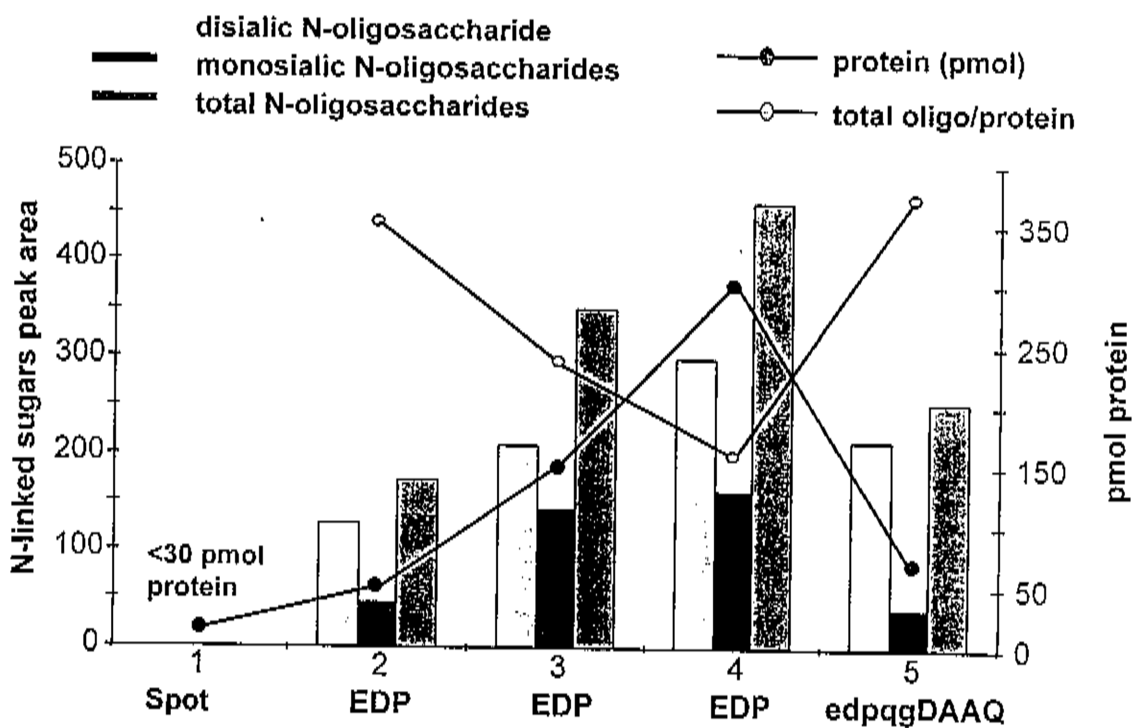
Steps involved in immunoassays:

Immune reaction (specific binding of the antibody with a target molecule)

Separation of antigen-antibody complexes from unreacted components

Visual or photometric detection of the reaction via (enzyme)-labelled reaction partners


Narrow pI Range PVDF blot of Human Plasma
4.5
5.5
200
α2HS
α1AT
1 2 3 4
1 2 3 4 5
M_r x 10^{-3}
8


Analysis of Human α1-Antitrypsin Isoforms
disialic N-oligosaccharide
monosialic N-oligosaccharides
total N-oligosaccharides
protein (pmol)
total oligo/protein
N-linked sugars peak area
pmol protein
500
400
300
200
100
0
350
250
150
50
0
<30 pmol
protein
1
Spot
2
EDP
3
EDP
4
EDP
5
edpqgDAAQ

Amino Acid Composition

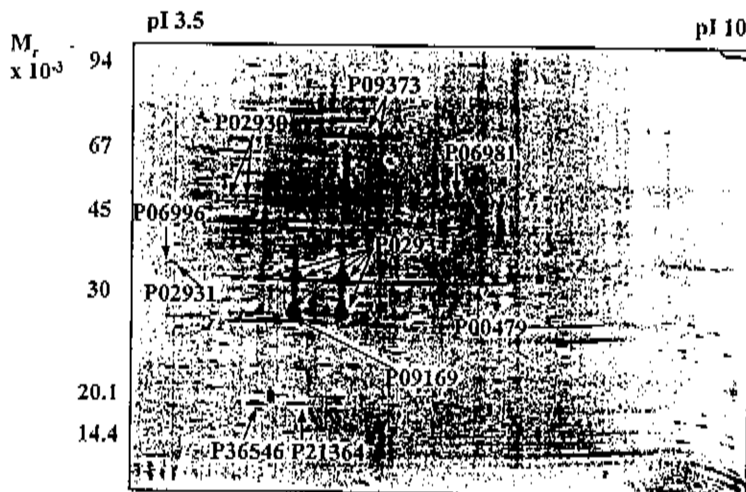
```
Spot ECOLItest
==============
AMINO ACID COMPOSITION
Asx: 17.90    Glx: 15.00    Ser:  3.40    Gly:  7.70
His: 0.30     Thr:  5.50    Ala: 14.10    Pro:  4.70
Tyr: 0.30     Arg:  1.40    Val:  9.30    Met:  0.70
Ile: 3.70     Leu:  4.50    Phe:  3.80    Lys:  8.50

Tagging: No_Tag
pI:    4.68   Range:  4.43  -  4.93
Mw:    9741   Range:  7793  - 11689

The ECOLI entries having pI and Mw values in the specified range:

Rank Score Protein      pI      Mw   Description
========================================================================================
   1    20 HDEA_ECOLI   4.68    9741  PROTEIN HDEA.
   2   187 GLR1_ECOLI   4.81    9685  GLUTAREDOXIN 1 (GRX1).
   3   187 YCCJ_ECOLI   4.70    8524  HYPOTHETICAL 8.5 KD PROTEIN IN AGP
   4   224 YFHF_ECOLI   4.43   11536  HYPOTHETICAL 11.5 KD PROTEIN IN HSCA
   5   224 THIO_ECOLI   4.67   11675  THIOREDOXIN.
```

Outer Membrane Proteins

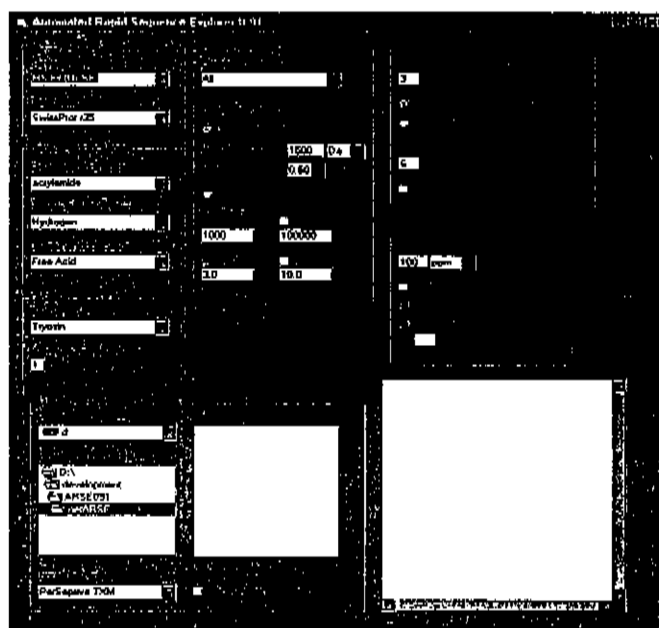


Swiss Prot Entry for *E. coli* protein

SWISS-PROT: P02339

ID SSB_ECOLI STANDARD; PRT; 177 AA.
AC P02339;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE SINGLE-STRAND BINDING PROTEIN (SSB) (HELIX-DESTABILIZING PROTEIN).
GN SSB OR EXRB OR LEXC.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-52.
RX MEDLINE; 82037821. [NCBI, ExPASy, Israel, Japan]
RA SANCAR A., WILLIAMS K.R., CHASE J.W., RUPP W.D.;
RL PROC. NATL. ACAD. SCI. U.S.A. 78:4274-4278(1981).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE; 94089392. [NCBI, ExPASy, Israel, Japan]
RA BLATTNER F.R., BURLAND V.D., PLUNKETT G. III, SOFIA H.J.,
RA DANIELS D.L.;
RL NUCLEIC ACIDS RES. 21:5408-5417(1993).

R.A.I.D.



Features

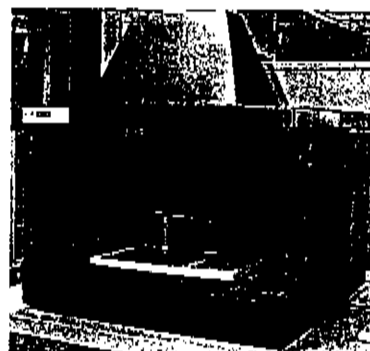
- Monoisotopic peak lists
- Filters contaminating masses (trypsin)
- Submits data to MS-Fit engine
- Dynamically adjusts user-definable parameters
- Uses the pI/Mr matching capabilities of MS-Fit
- Extracts pI/Mr information from Melanie image

The ARRM-214

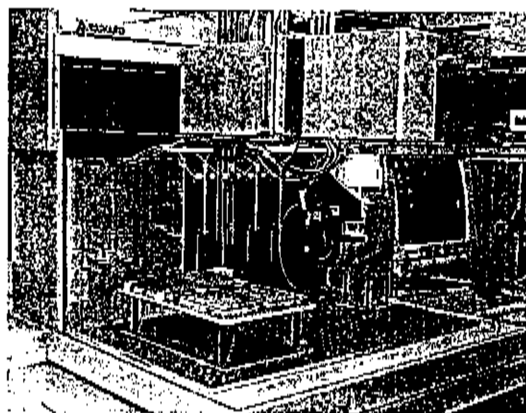
Features

- Fast
- No mistakes
- No contamination

- CCD camera system
- Precision x-y arm
- "Point and click" with data logging
- Delivers to microtitre plate



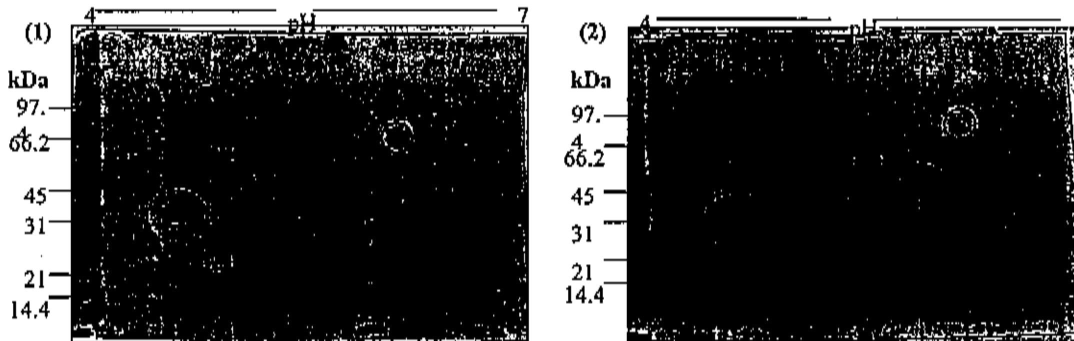
Automated Digestion with Multiprobe 104



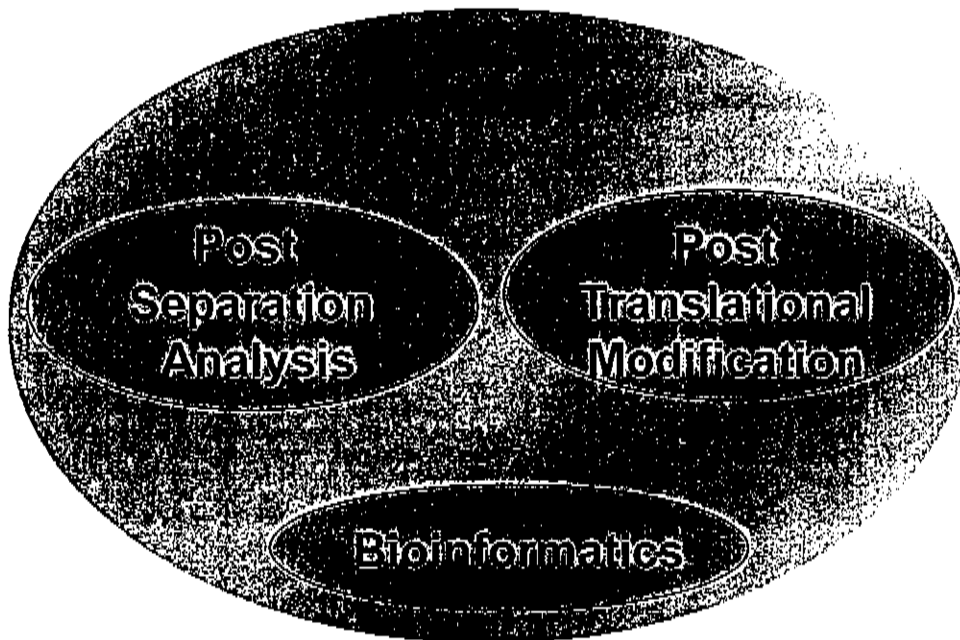
Features

- Fast and error free
- Adapted to perform all steps
- Automatically delivers matrix and peptide containing supernatant to a MALDI-TOF flat 100 sample stage

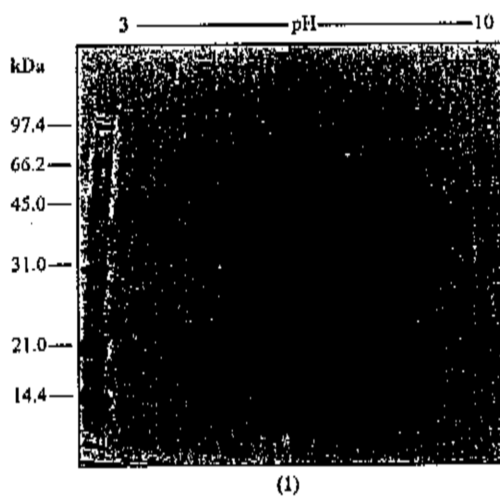
Differential Display of Wheat Proteins



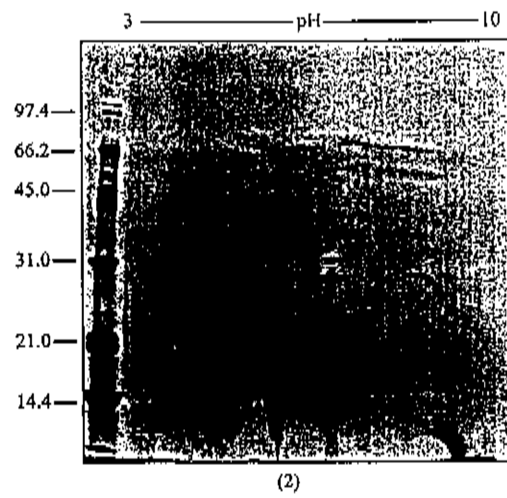
Characterisation



Sequential Extract of Wheat

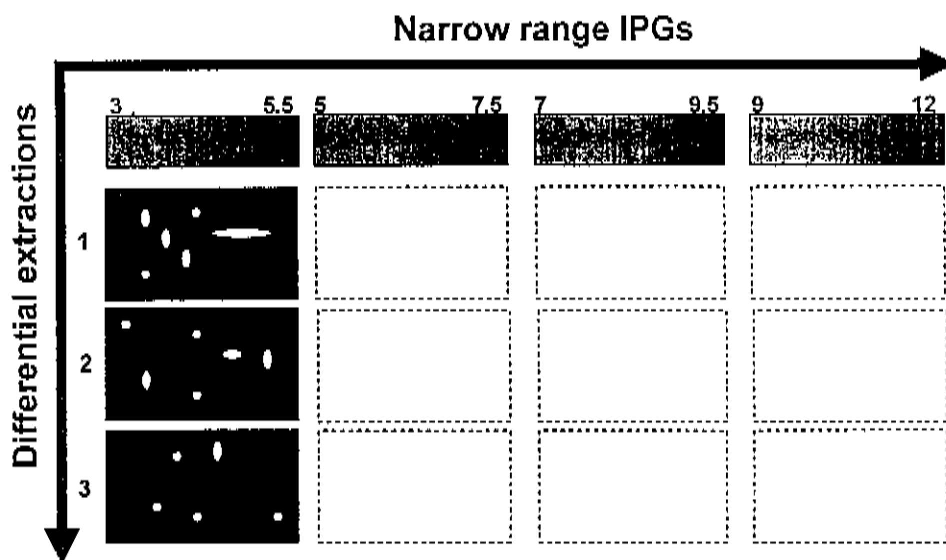


(1)

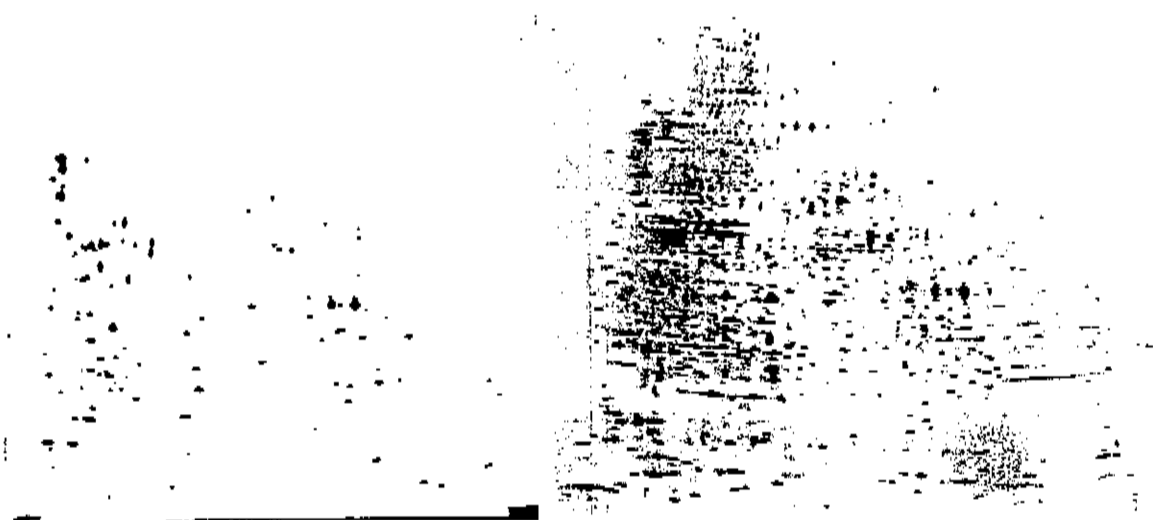


(2)

Arraying complex mixtures- seeing the low abundance proteins



Improved Solubility with Thiourea and Sulfobetaine Surfactants

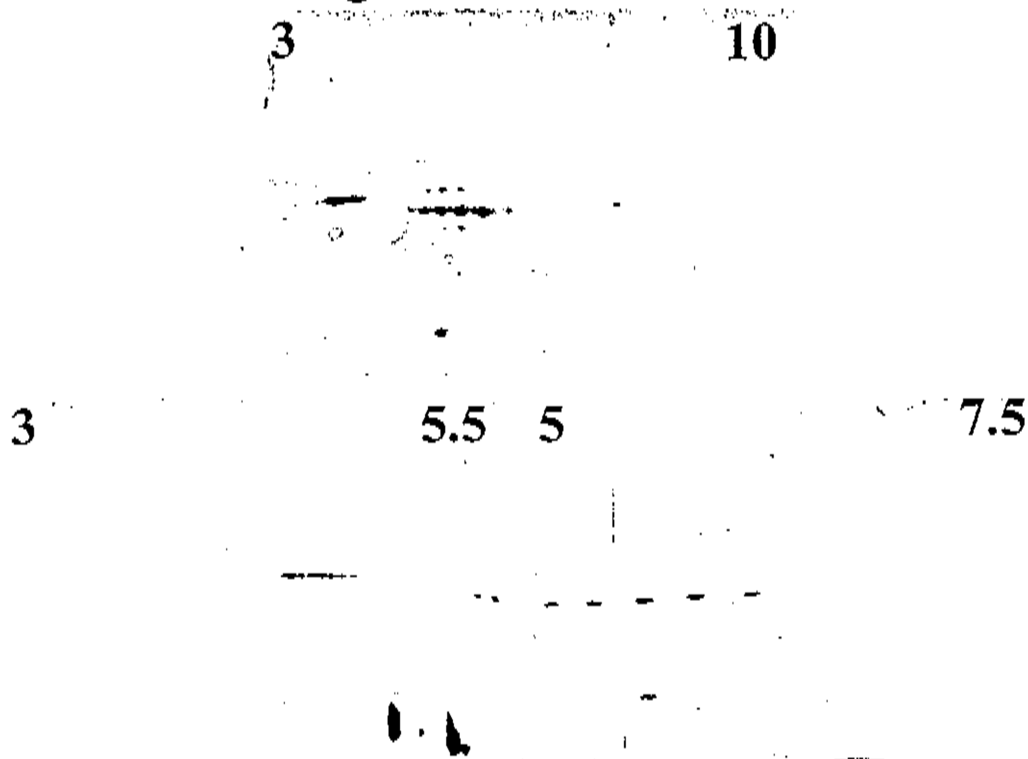


Conventional - Urea, CHAPS, DTT

Improved - Urea, Thiourea, CHAPS, SB 3-10, TBP

100μg loads of *E. coli*

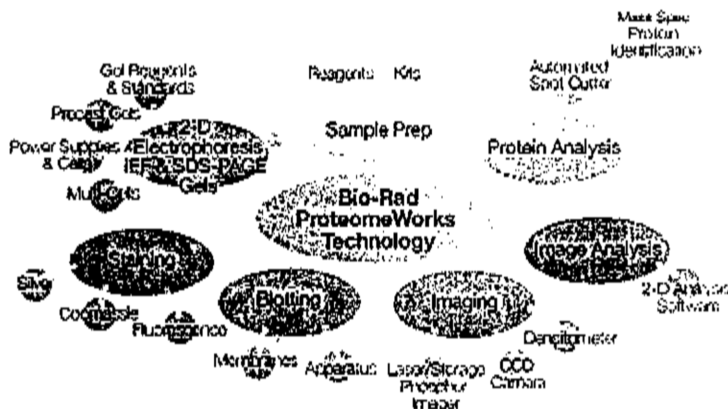
Narrow Range IPGs Increase Resolution



What do these places have in common ?

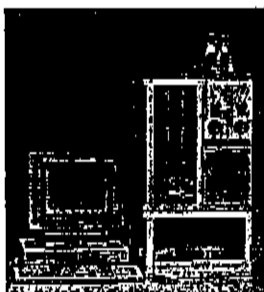
www.apbiotech.com

Committed to a complete line of proteomics products



Traditional Protein Chemistry

- Handle proteins one at a time



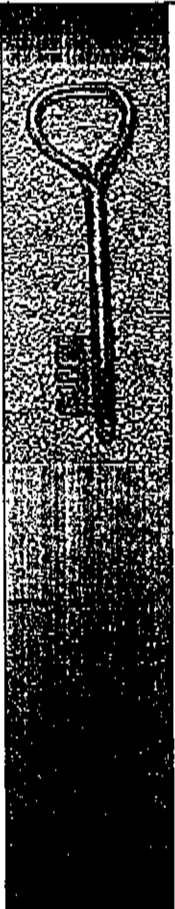
Proteomics

- Complex samples arrayed
- Don't purify one protein, purify them all

Proteomics

Putting it all together

BRAD WALSH, APAF



Plant BreedingCompanies

PIONEER HI-BRED INTERNATIONAL, INC.

Rahman, S., Abrahams, S., Abbott, D., Mukai, Y., Samuel, M., Morell, M., Appels, R. 1997. A complex arrangement of genes at a starch branching enzyme I locus in the D-genome donor of wheat. Genome 40, 465 - 474.

Rahman, S., Jolly, CJ., Skerritt, JH., Walloshcck, A. 1994. Cloning of a wheat 15-kDa grain softness protein (GSP). GSP is a mixture of puroindoline-like polypeptides. Eur. J. Biochem. 223, 917 - 925.

Rasmussen,S.K., Norgaard,A., Hejgaard,J. and Dahl,S.W. 1996. A recombinant wheat serpin with inhibitory activity. Plant Mol. Biol. 30, 673-677.

Reddy, P., Appels, R. 1992. Analysis of a genomic DNA segment carrying the wheat high molecular weight (HMW) glutenin Bx17 subunit and its use as a RFLP marker. Theor. Appl. Genet.

Repellin, A., Nair, RB., Bagar, M., Chibbar, RN. 1997. Isolation of starch branching enzyme I cDNA from a wheat endosperm library. Plant gene register, accession # Y12320 (www.tarweed.com/pgr/PGR 97-094.html)

Rounsley, SD., Glodek, A., Sutton, G., Adams, MD., Somerville, CR., Venter, JC., Kerlavage, AR. 1996. The construction of *Arabidopsis* expressed sequence tags assemblies. Plant Physiol. 112, 1177 - 1183.

Saski, T., Song, J., Koga-Ban, Y., Matsui, E., Fang, F., Higo, H., Nagasaki, H., Hori, M., Miya, M., Murayama-Kayano, E., Takiguchi, T., Takasuga, A., Niki, T., Ishimaru, K., Ikeda, H., Yamamoto, Y., Mukai, Y., Ohta, I., Miyadera, N., Havukkala, I., Minobe, Y. 1994. Toward cataloguing all rice genes: large-scale sequencing of randomly chosen rice cDNAs from a callus cDNA library. The Plant J. 6, 615 - 624.

Schena, M., Shalon, D., Davis, RW., Brown, PO. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science, 270, 467 - 470.

Sheridan, WF., Clark, JK. 1993. Mutational analysis of morphogenesis of the maize embryo. The Plant J. 3, 347 - 358.

Shewry, PR., Miles, MJ., Tatham, A. 1994. The prolamin storage proteins of wheat and related cereals. Prog. Biophys. Molec. Biol. 61, 37 - 59.

Shimoni, Y., Zhu, X-Z., Levanony, H., Degal, G., Galili, G. 1995. Purification, characterisation, and intracellular localisation of glycosylated protein disulfide isomerase from wheat grains. Plant Physiol. 108, 327 - 335.

Shimoni,Y., Zhu,X.Z., Segal, G. and Galili,G. 1994. Genbank U11496

Singh,N.K., Donovan,G.R., Carpenter,H.C., Skerritt,J.H., Langridge,P. Isolation and characterization of wheat triticin cDNA revealing a unique lysine-rich repetitive domain. Plant Mol. Biol. 22 (2), 227-237 (1993)

Skerritt JH., Lew PY., Castle SL 1988. Accumulation of gliadin and glutenin polypeptides during development of normal and sulphur-deficient wheat seed: analysis using specific monoclonal antibodies. Journal Experimental Botany. 39, 723-737

Smart MG., O'Brien TP. 1983. The development of the wheat embryo in relation to the neighbouring tissues. Protoplasma, 114: 1-13

Smith,J.J. and Raikhel,N.V. 1989. Nucleotide sequences of cDNA clones encoding wheat germ agglutinin isolectins A and D. Plant Mol. Biol. 13, 601-603.

Soares, MB., Bonaldo, MdeF., Jelene, P., Su, L., Lawton, L., Efstradiadis, A. 1994. Construction and characterisation of a normalised cDNA library. 1994. Proc. Natl. Acad. Sci. USA. 91, 9228 - 9232.

Spielmeyer, W., Robertson, M., Collins, N., Leister, D., Schulze-Lefert, P., Seah, S., Moullet, O., Lagudah, ES. A superfamily of disease resistance gene analogs is located on all homoeologous chromosome groups of wheat (*Triticum aestivum*). Genome, in press

Tabata,T., Nakayama,T., Mikami,K. and Iwabuchi,M. 1991. HBP-1a and HBP-1b: leucine zipper-type transcription factors of wheat. EMBO J. 10 (6), 1459-1467.

Tabata,T., Takase,H., Takayama,S., Mikami,K., Nakatsuka,A., Kawata,T., Nakayama,T. and Iwabuchi,M. 1989. A protein that binds to a cis-acting element of wheat histone genes has a leucine zipper motif. Science 245, 965-967

Tabata,T., Terayama,C., Mikami,K., Uchiyama,H. and Iwabuchi,M. 1987. An accurate transcription of wheat histone genes in sunflower cells. Plant Cell Physiol. 28, 73-82

Ugalde TD., Jenner CF. 1990. Route of substrate movement into wheat endosperm. I. Carbohydrates. Australian Journal of Plant Physiology. 17, 693-704.

Wagner, GB., Haeger, KP., Ziegler, P. 1996. Nucleotide sequence of a cDNA from wheat leaves encoding ubiquitous β-amylase. Gen Bank TAAMY1

Wang HL., Patrick JW., Offler CE., Wang XD. 1995. The cellar pathway of photosynthate transfer in the developing wheat grain. III. A structural analysis and physiological studies of the pathway from the endosperm cavity to the starchy endosperm. Plant, Cell and Environment. 18, 389-407

Wang N., Fisher DB. 1995. Sucrose release into the endosperm cavity of wheat grains apparently occurs by facilitated diffusion across the nucellar cell membranes. Plant Physiology. 109, 579-585

Wang, TL., Bogracheva, TY., Hedley, CL. 1998. Starch: as simple as A, B, C?. J. Exp. Botany 49, 481 - 502.

Wang,TL., Cuming, A. 1995. Embryogenesis: the generation of a plant. BIOS scientific publishers.

Weber, H., Borisjuk, L., Heim, U., Sauer, N., Wobus, U. 1997. A role for sugar transporters during seed development: molecular characterization of a hexose and a sucrose carrier in Faba bean seeds. The Plant Cell 9, 895 - 908.

Wilkins, MR., Williams, KL. 1998. Discovering databases. Today's Life Science, April, p26 - 31.

Yamamoto, K., Sasaki, T. 1997. Large-scale EST sequencing in rice. Plant Mol. Biol. 35, 135 - 144.

Yan, L., Bhave, M., Fairclough, R., Konik, C., Rahman, S., Appels, R. 1998. The genes encoding granule-bound starch synthase (GBSS) from the A,B, and D progenitors of common wheat. Plant Syst. Evol. Submitted.

Zee, S.-Y., O'Brien, TP. 1970. Studies on the ontogeny of the pigment strand in the caryopsis of wheat. Australian J. Biol. Sci. 23, 1153 - 1171.

References

Ainsworth, CC., Clark, J., Balsdon, J. 1993. Expression, organisation and structure of the genes encoding the *Waxy* protein (granule bound starch synthase) in wheat. Plant Mol. Biol. 22, 67 - 82.

Albani, D., Hammond-Kosack, MCU., Smith, C., Conlan, S., Colot, V., Holdsworth, M., Bevan, MW. 1997. 9, The Plant Cell 171 - 184.

Alrefai, R. Berke, TG, Rocheford, TR, 1995. Quantitiative trati locus analysis of fatty acid concentrations in maize. Genome 38,894-901.

Anderson, O. 1996. *Triticum aestivum* a-gliadin. GenBank TAU50984

Andrade, MA., Sander, C. 1997. Bioinformatics: from genome data to biological knowledge. Current Opinion in Biotechnology. 8, 675 - 683.

Apsit,V., Freeberg,J.A., Chase,M.R., Davis,E.A. and Ackerman,S.J. Wheat TFIID TATA binding protein Nucleic Acids Res. 21, 1494-1494 (1993)

Benson, DA., Boguski, MS., Lipman, DJ., Ostell, J., Ouellette, BFF. 1998. GenBank. Nucleic Acids Res. 26, 1 - 7.

Blumenthal, C., Rahman, S., Howarth, C., Appels, R. The HS70 protein genes in wheat. Manuscript in preparation

Bonaldo, M.deF., Lennon, G., Soares, M.B. 1996. Normalization and subtraction: two approaches to facilitate gene discovery. Genome research 6, 791 - 806.

Briarity LG., Hughes, CE., Evers AD .1979. The development endosperm of wheat a stereological analysis. Annals of Botany, 44, 641-658

Castagnaro,A., Marana,C., Carbonero,P. and Garcia-Olmedo,F. 1994. cDNA cloning and nucleotide sequences of alpha 1 and alpha 2 thionins from hexaploid wheat endosperm. Plant Physiol. 106 (3), 1221-1222.

Chandler, PM., Robertson, M. 1994. Gene expression regulated by abscisic acid and its relation to stress tolerance. Annu. Rev. Plant Physiol. Plant Mol. Biol. 45, 113 - 141.

Ciaffi, M., Lee, YK., Tamas, L., Gupta, R., Skerritt, J., Appels, R. 1998. Molecular characterisation of LMW glutenin subunit genes in the D genome donor of hexaploid wheat (*Triticum aestivum*). Theor. Appl. Genet. In press.

Clarke, BC., Appels, R. 1998. Sequence variation at the *Sec-1* locus of rye. Plant Syst. Evol., in press

Cooke, R., Raynal, M., Laudie, M., Grellet, F., Delseny, M., Morris, P-C., Guerrier, D., Giraudat, J., Quigley, F., Clabault, G., Li, Y-F., Mache, R., Krvitzky, M., Gy, IJ-J., Kreis, M., Lecharny, A., Parmentier, Y., Marbach, J., Fleck, J., Clement, B., Philipps, G., Herve, C., Bardet, C., Tremousaygue, D., Lescure, B., Lacomme, C., Roby, D., Jourjon, M-F., Chabrier, P., Charpenteau, J-L., Desprez, T., Amselem, J., Chiapello, H., Hofte, H. 1996. Further progress towards a catalogue of all *Arabidopsis* genes: analysis of a set of 500 non-redundant ESTs. The Plant J. 9, 101 - 124.

Curry,J., Morris,C.F. and Walker-Simmons,M.K. 1991. Sequence analysis of a cDNA encoding a group 3 LEA mRNA inducible by ABA or dehydration stress in wheat. Plant Mol. Biol. 16 (6), 1073-1076.

Danyluk,J., Houde,M., Rassart,E. and Sarhan,F. 1994. Differential expression of a gene encoding an acidic dehydrin in chilling sensitive and freezing tolerant gramineae species. FEBS Lett. 344, 20-24.

Doan, DNP., Linnestad, C., Olsen, O-A. 1996. Isolation of molecular markers from the barley endosperm coenocyte and the surrounding nucellus layers. Plant Mol. Biol. 31, 877 - 886.

Dratewka-Kos, E., Rahman, S., Grzekzak, ZF., Kennedy, TD., Murray, RK., Lane, BG. 1989. The polypeptide structure of germin as deduced from cDNA sequencing. J. Biol. Chem. 264, 4896 - 4900.

Feuillet, C., Schachermayr, G., Keller, B. 1997. Molecular cloning of a new receptor-like kinase gene encoded at the *Lr10* disease resistance locus of wheat. The Plant J. 11, 45 - 52.

Feuillet,C., Messmer,M., Schachermayr,G. , Keller,B. Genetic and physical characterization of the LR1 leaf rust resistance locus in wheat (Triticum aestivum L.) Mol. Gen. Genet. 248 (5), 553-562 (1995)

Gao, XP., Francis D., Ormrod JC., Bennett MD. 1992 . Changes in cell number and cell division activity during endosperm development in allohexaploid wheat Triticum aestivum L.. Journal of Experimental Botany. 43, 1603-1609.

Garcia-Marolo, F., Marana, C., Mena, M., Garcia-Olmedo, F., Carbonero, P. 1990. Cloning of cDNA and chromosomal location of genes encoding the three types of subunits of the wheat tetrameric inhibitor of α-amylase. Plant Mol. Biol. 14, 845 - 853.

Gautier,M.F., Alary,R. and Joudrier,P. 1990. Cloning and characterization of a cDNA encoding the wheat (Triticum durum Desf.) CM16 protein. Plant Mol. Biol. 14 (3), 313-322.

Gautier, MF, Aleman, M-E., Guirao, A., Marion, D., Joudrier, P. 1994. Triticum aestivum puroindolines, two basic cystine-rich seed proteins: cDNA analysis and developmental gene expression. Plant Mol. Biol. 25, 43 - 57.

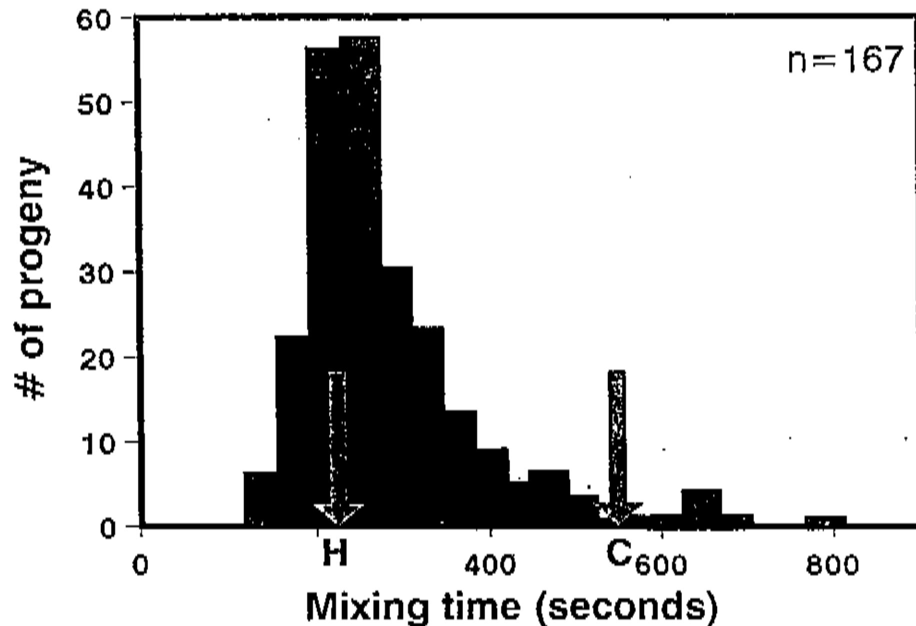
Grimwade, B., Tatham, A., Fredman, RB., Shewry, PR., Napier, JA. 1996. Comparison of the expression patterns of genes coding for wheat gluten proteins and proteins involved in the secretory pathway in developing caryopses of wheat. Plant Mol. Biol. 30, 1067 - 1073.

Gubler, F., Kalla, R., Roberts, TK., Jacobsen, JV. 1995. Gibberellin regulated expression of a Myb gene gene in barley aleurone cells: Evidence for Myb transactivation of a high-pI α-amylase gene promoter. Plant Cell 7, 1879 - 1891.

Sequence database for wheat genes

Sequence catagory	number sequenced	selected references
ITMI mapped cDNA's	734	Lazo, et al (1998)
cDNA's from wheat endosperm library (not normalised)	500	Lazo and Anderson (unpub.)
chaperones and transcription factors (cDNAs and genomic clones):		
protein disulfide isomerase		Napier et al (1994)
		Shimoni et al (1994)
basic leucine zipper protein (bZIP)	5	Albani et al (1997)
		Tabata et al (1989, 1991)
		Niu and Guiltinan, 1994)
		Mikami et al (1994)
lumenal binding protein (BiP)		Grimwade et al (1996)
heat shock proteins (HS18, HS70)	20	Blumenthal et al (1998)
		Joshi and Nguyen (1994)
GAMyb (barley)		Gubler et al (1995)
transcription factor TFIID		Apsit et al (1993)
starch biosynthetic/degrading enzymes (cDNAs and genomic clones):		
sucrose synthase		Marana et al (1988)
ADP-glucose pyrophosphorylase		Olive et al (1989)
granule bound starch synthase (GBSS)		Ainsworth et al (1993)
		Yan et al (1998)
starch branching enzyme I		Rahman et al (1997)
		Repellin et al (1997)
starch branching enzyme II		Nair et al (1997)
		Rahman et al (1998)
soluble starch synthase I		Li et al (1998)
α-amylase		Huttly et al 1988)
β-amylase		Wagner et al (1996)
genes linked to hardness (*Ha*) locus (cDNAs and genomic clones):		
grain softness protein		Rahman et al (1994)
puroindolin a and b		Gautier et al (1994)
grain protein genes (cDNAs and genomic clones):		
high molecular weight glutenin subunit proteins	42	Reddy and Appels (1992)
		Shewry et al, (1994)
low molecular weight glutenin subunit proteins	15	Shewry et al (1994)
		Ciaffi et al, 1998)
gliadins	46	Shewry et al (1994)
		Anderson (1996)
serpin		Rasmussen et al (1996)
agglutinin isolectins		Smith and Raikhel (1989)
purothionin/thionin	4	Castagnaro et al (1994)
α-amylase inhibitor	4	Garcia-Marolo et al (1990)
β 1,3-glucanase	4	Fincher (1995)
CM-protein	3	Gautier et al (1990)
triticin		Singh et al (1993)
7S globulin		Litts et al (1991)
disease resistance gene (cDNAs and genomic clones):		
Cre3 locus on chromosome 2D		Lagudah et al (1997)
Lr1 locus on chromosome		Feuillet et al (1995)
Lr10 locus on chromosome 1A		Feuillet et al (1997)
resistance gene analogues (RGAs)		Spielmeyer et al (1998)
cold response proteins	10	Danyluk et al (1994)
		Oullet et al (1998)
chromosomal proteins - histones	18	Tabata et al (1987)
chloroplast/mitochondrial proteins	10	www.ncbi.nlm.nih.gov/
"household" proteins/enzymes	120	www.ncbi.nlm.nih.gov/

Analysis of progeny from Cranbrook x Halberd cross



ra98857a-4

Gene discovery in the area of wheat quality

- **Interpretation of data**
  - micro-arrays to identify genes expressed at critical stages of development,
  - mutational analyses to identify genes of interest. Utilise synteny among grass genomes,
  - mapping/chromosome walking capabilities,
  - a clear definition of the phenotypes that define quality attributes.

Gene discovery in the area of wheat quality

- BAC clones are a major tool for chromosome walking

  - several regions of wheat chromosomes are of specific interes in relation to quality

  - distal region of chromosome 1 (gliadin, LMW glutenins)

  - distal region of 5D (hardness locus)

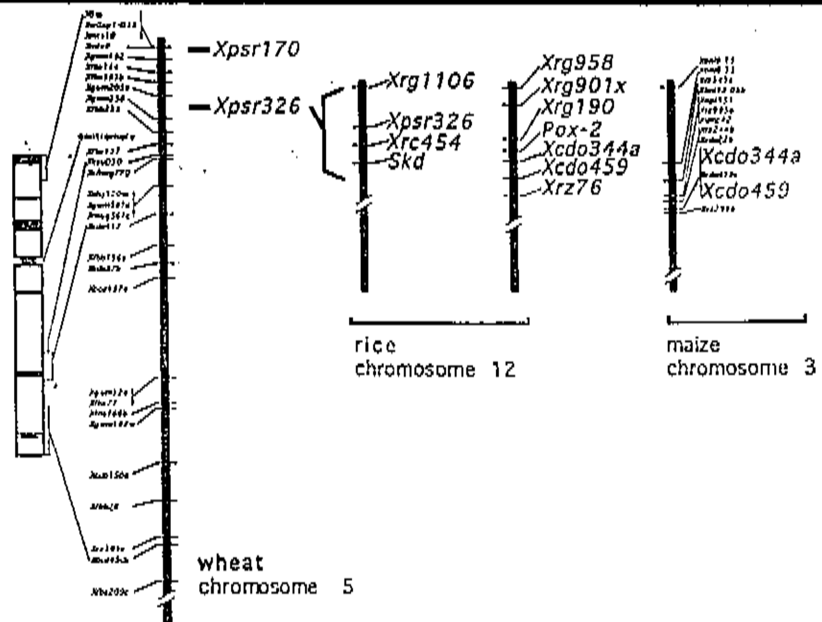
  - short arm of chromosome 7 (starch biosynthetic enzymes)

BAC clones (from the D genome) are being characterised for each of these regions. The levels of genetic recombination are such that significant genetic distances are covered by clones 100kb in length.

CSIRO PLANT INDUSTRY

Comparative analysis of grass genomes

Tracing a wheat chromosome segment in rice and maize



RA98645-2

CSIRO

Gene discovery in the area of wheat quality

- Large-scale sequencing of expressed genes

Other genes identified (alignments unambiguous) that are of potential interest to determining wheat flour quality:

Protein disulfide isomerase (4)

Peptidyl prolyl isomerase (3)

HSP70 (2)

Calnexin (1)

Gene discovery in the area of wheat quality

- Large-scale sequencing of expressed genes

The grid of 25,000 clones is also being screened for families of genes that we predict will be important for determining wheat quality. For example:

nutrient transporters

transcription factors

These families of genes then go into the sequencing effort to contribute to the data base.

Gene discovery in the area of wheat quality

- Large-scale sequencing of expressed genes

The fundamental outputs from the search are HSPs (high scoring segment pairs)

The score is recorded in E values (at low value this approximates the probability of the match having occurred by chance):

$E = k N \exp(-\lambda S)$

N = query sequence length x data base sequence length
λ = reliability of alignment
s= number of matches

Gene discovery in the area of wheat quality

- Large-scale sequencing of expressed genes

35% of sequences had matches with probability scores <E - 10 (equivalent to $P<10^{-10}$)

10.5% unknown

8% ribosomal proteins genes
6.7% seed storage protein genes (HMW glutenins, 0.9%, LMW glutenins, 1.3%, gliadins 4.5%)
2.2% histones
2% elongation factor proteins

CSIRO

Gene discovery in the area of wheat quality

- Establish cDNA and BAC libraries
- Large-scale sequencing of expressed genes
- Bioinformatics (storage and analysis of data)
- Interpretation of structural information

Gene discovery in the area of wheat quality

- Establish cDNA and BAC libraries

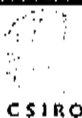
A normalised cDNA library has been established for barley endosperm development. The normalisation procedure reduces the number of copies of any one gene in the mixture and thus makes random sequencing more efficient.

A normalised wheat cDNA library is being established and a very good standard one is already available in a directional cloning vector (Xho1/EcoR1). Several libraries exist overseas as well.

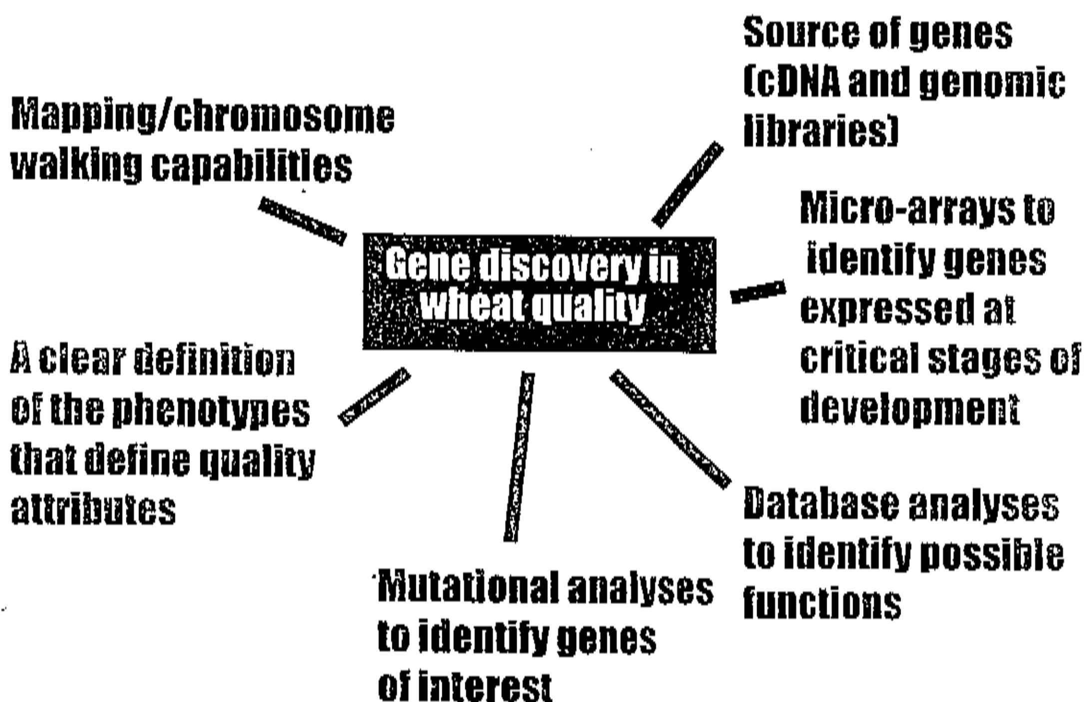
A 4x coverage of the D genome of wheat in BAC clones (80-150 kb insert lengths) has been produced (O. Moullet, CSIRO-PI)

Gene discovery in the area of wheat quality

- The structure of genes controlling product quality
- DNA sequences controlling the expression of "quality" genes
- An intellectual property position on genes determining key features of the quality of Australian agricultural products



Gene discovery in the area of wheat quality



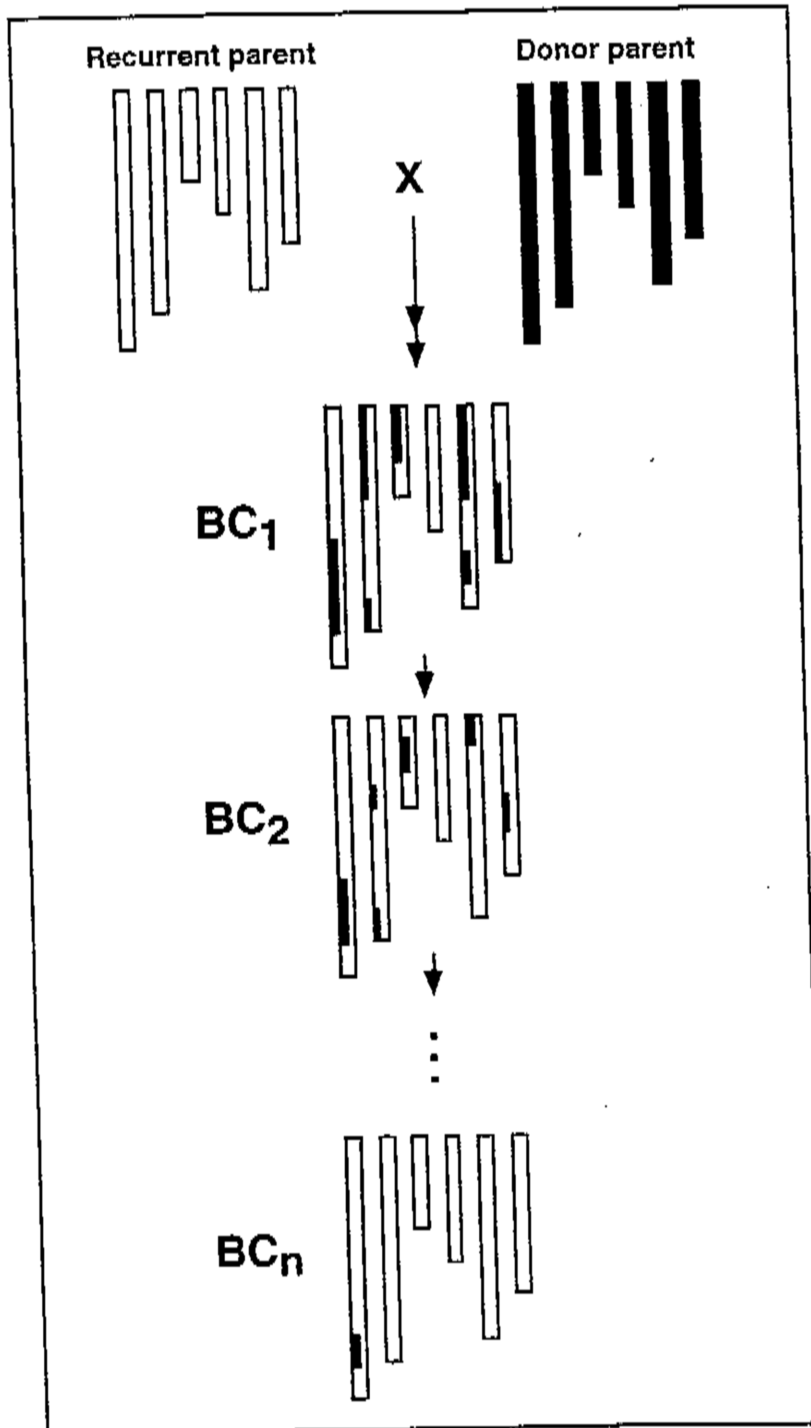
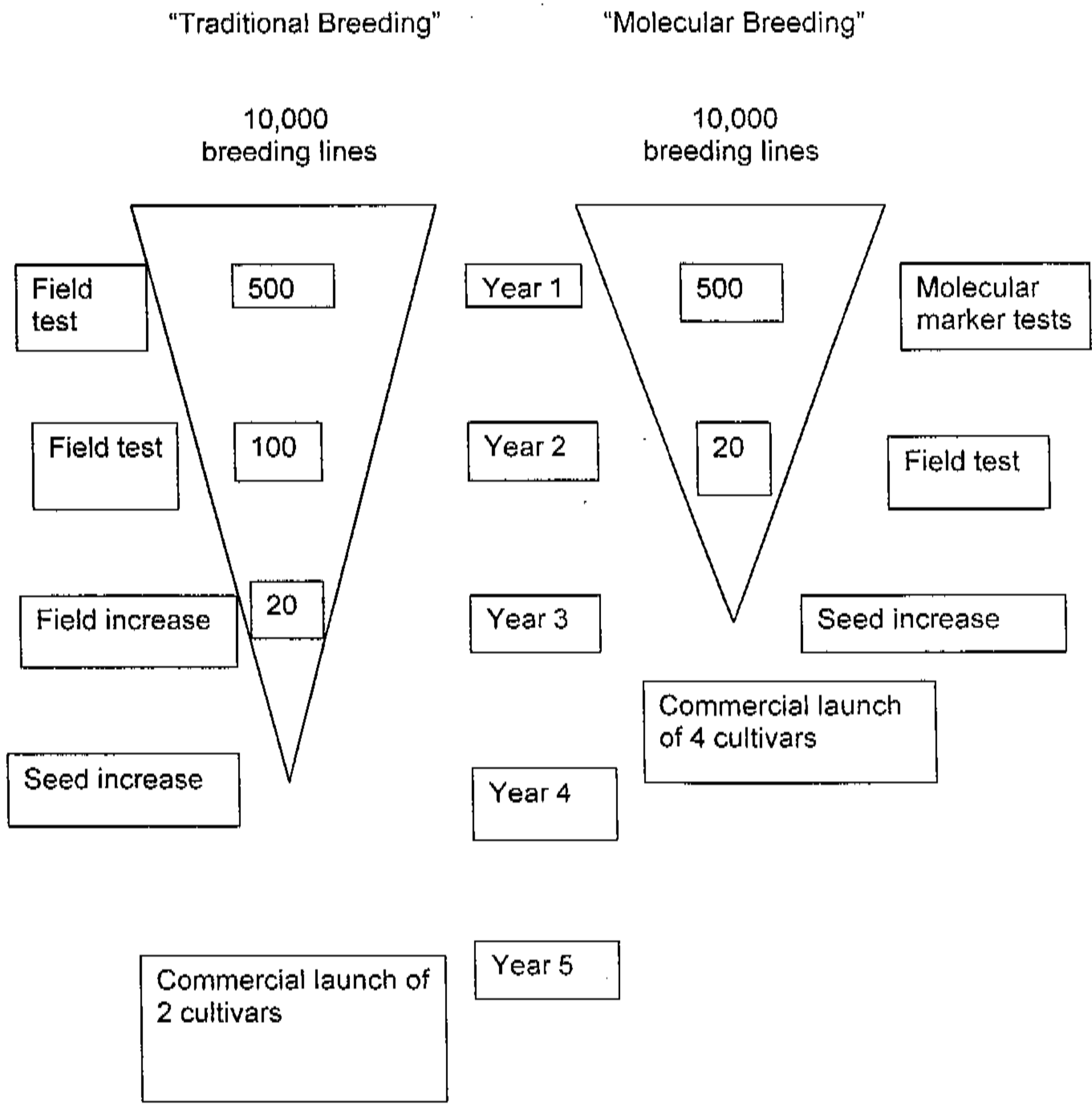


Fig.1. *Development of NILs by backcrossing. Schematic demonstration of the introgression of a genomic segment from the donor into the recurrent parent.*

GBSS gene microsatellite confirmation of F1 plants

DHWx12 Goldmark — confirmed F1 plants —

AIM OF MOLECULAR BREEDING:



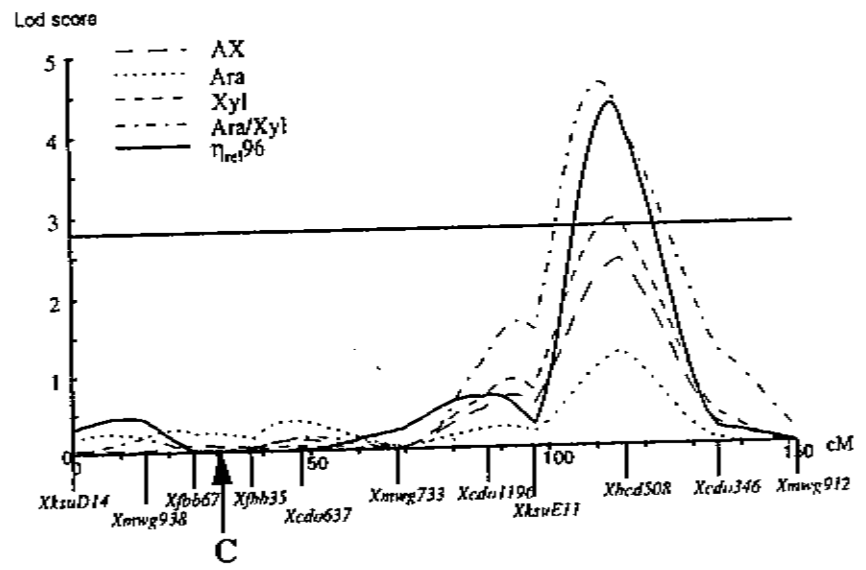


Fig. 2 Representation of the major QTL located on 1BL chromosome for the ITMImap cross. *C* Centromere

TAGGING?

1. Complete mapping and linkage analysis

- Single genes, sometimes with candidate gene approach, and sometimes with knowledge of position of gene.

- QTL approach (quantitative trait locus)
  Requires good map, and statistical approach to associating trait and markers.

2. Bulked segregant analysis

- Single genes, or "strong" QTL(?).

- Best when using "multilocus" technique (RAPDs, AFLPs), unless there is some knowledge of the position of the gene.

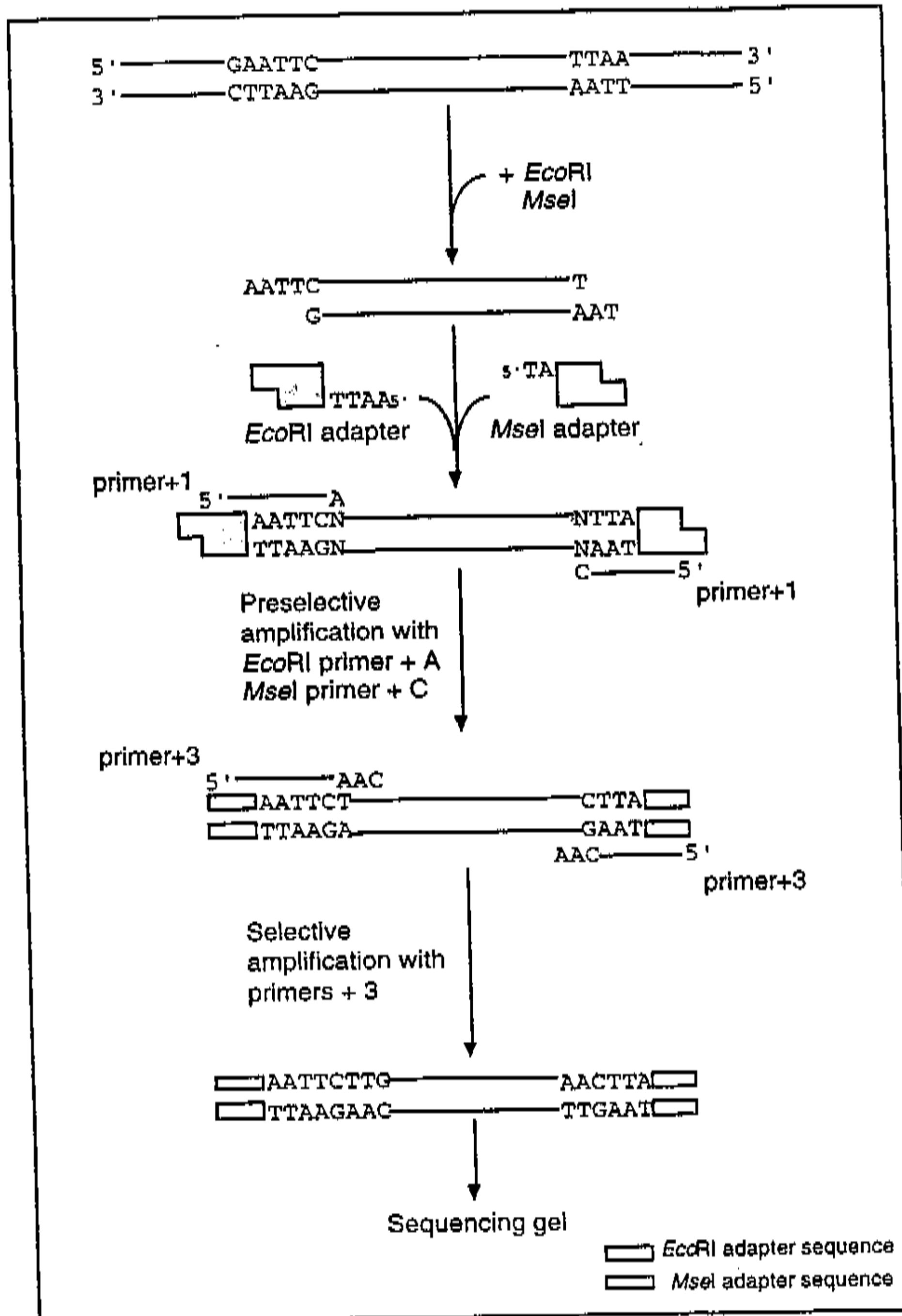


Fig. 3 The AFLP procedure using two primer pairs.

Genomic DNA is digested by restriction endonucleases EcoRI and MseI, ligated to their respective adapters and selectively amplified using primers that contain the sequences of the adapters and one to three arbitrary nucleotides as selective sequences. Polymorphisms of these selectively amplified segments are resolved in a sequencing gel.

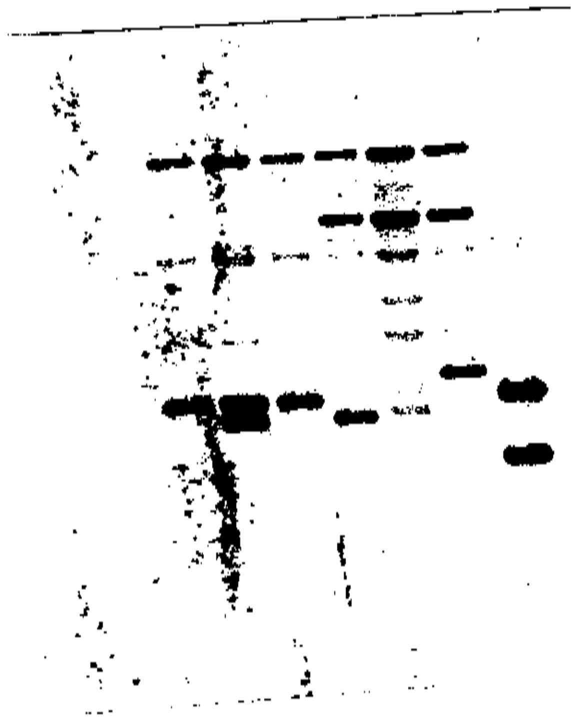
APPENDIX

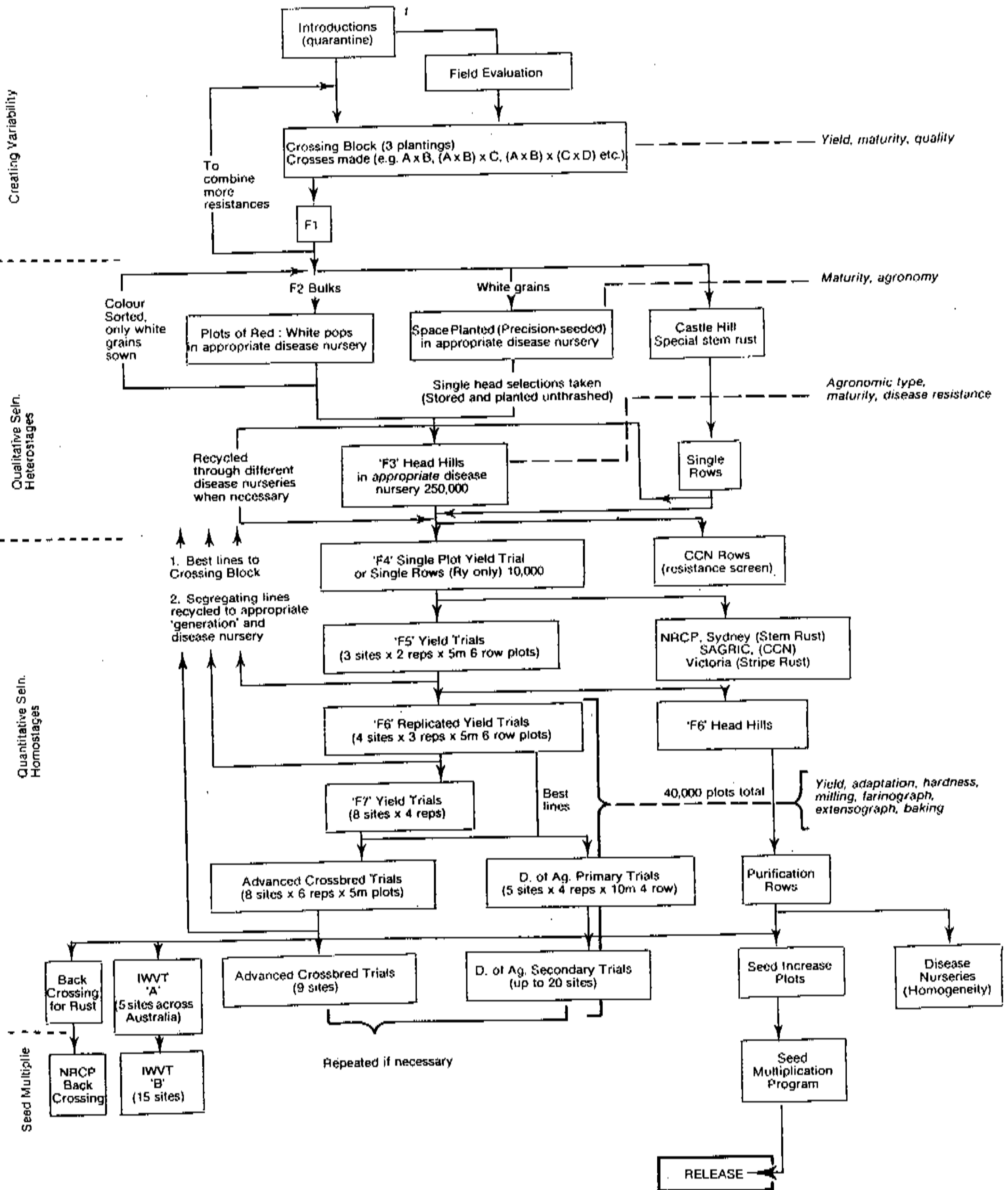
Description of wheat microsatellite primer sets and loci

Locus	Left primer	Right primer	Repeat	An. temp.	Opata (bp)	Synth. (bp)
Xgwm2-3A	CTG CAA GCC TGT GAT CAA CT	CAT TCT CAA ATG ATC GAA CA	(CA)18	50°	128	130
Xgwm2-3D	CTG CAA GCC TGT GAT CAA CT	CAT TCT CAA ATG ATC GAA CA	(CA)18	50°	265	267
Xgwm3-3D	GCA GCG GCA CTG GTA CAT TT	AAT ATC GCA TCA CTA TCC CA	(CA)18	55°	84	—
Xgwm4-4A	GCT GAT GCA TAT AAT GCT GT	CAC TGT CTG TAT CAC TCT GCT	(CA)13(TA)26	55°	257	255
Xgwm5-3A	GCC AGC TAC CTC GAT ACA ACT C	AGA AAG GGC CAG GCT AGT AGT	(TC)23(T)4(GT)12(GA)10	50°	171	158
Xgwm6-4B	CGT ATC ACC TCC TAG CTA AAC TAG	AGC CTT ATC ATG ACC CTA CCT T	(GA)40	55°	207	196
Xgwm10-2A	CGC ACC ATC TGT ATC ATT CTG	TGG TCG TAC CAA AGT ATA CGG	(AT)5(GT)15	50°	138	143
Xgwm11-1B	GGA TAG TCA GAC AAT TCT TGT G	GTG AAT TGT GTC TTG TAT GCT TCC	(TA)6CATA(CA)19(TA)6	50°	202	213
Xgwm16-2B	GCT TGG ACT AGC TAG AGT ATC ATA C	CAA TCT TCA ATT CTG TCG CAC GG	(C)12ACAAA(CA)14(GA)18	50°	181	176
Xgwm16-5D	GCT TGG ACT AGC TAG AGT ATC ATA C	CAA TCT TCA ATT CTG TCG CAC GG	(C)12ACAAA(CA)14(GA)18	50°	224	225
Xgwm16-7B	GCT TGG ACT AGC TAG AGT ATC ATA C	CAA TCT TCA ATT CTG TCG CAC GG	(C)12ACAAA(CA)14(GA)18	50°	206	204
Xgwm18-1B	TGG CGC CAT GAT TGC ATT ATC TTC	GGT TGC TGA AGA ACC TTA TTT AGG	(CA)17GA(TA)4	50°	188	182
Xgwm30-2D	ATC TTA GCA TAG AAG GGA GTG GG	TTC TGC ACC CTG GGT GAT	(AT)19(GT)15	60°	—	156
Xgwm30-3A	ATC TTA GCA TAG AAG GGA GTG GG	TTC TGC ACC CTG GGT GAT	(AT)19(GT)15	60°	196	205
Xgwm32-3A	TAT GCC GAA TTT GTG GAC AA	TGC TTG GTC TTG AGC ATC AC	(GA)19	55°	169	173
Xgwm33-1A	GGA GTC ACA CTT GTT TGT GCA	CAC TGC ACA CCT AAC TAC CTG C	(GA)19	60°	116	—
Xgwm33-1B	GGA GTC ACA CTT GTT TGT GCA	CAC TGC ACA CCT AAC TAC CTG C	(GA)19	60°	—	119
Xgwm33-1D	GGA GTC ACA CTT GTT TGT GCA	CAC TGC ACA CCT AAC TAC CTG C	(GA)19	60°	—	158
Xgwm37-7D	ACT TCA TTG TTG ATC TTG CAT G	CGA CGA ATT CCC AGC TAA AC	(AG)8CG(AG)21	60°	189	—
Xgwm43-7B	CAC CGA CGG TTT CCC TAG AGT	GGT GAG TGC AAA TGT CAT GTG	(CA)22	60°	184	176
Xgwm44-7D	GTT GAG CTT TTC AGT TCG GC	ACT GGC ATC CAC TGA GCT G	(GA)28	60°	178	176
Xgwm46-7B	GCA CGT GAA TGG ATT GGA C	TGA CCC AAT AGT GGT GGT CA	(GA)2GC(CA)33	60°	186	179
Xgwm47.1-2A	TTG CTA CCA TGC ATG ACC AT	TTC ACC TCG ATT GAG GTC CT	(CT)7TT(CT)16	60°	—	170
Xgwm47.2-2A	TTG CTA CCA TGC ATG ACC AT	TTC ACC TCG ATT GAG GTC CT	(CT)7TT(CT)16	60°	150	—
Xgwm47-2B	TTG CTA CCA TGC ATG ACC AT	TTC ACC TCG ATT GAG GTC CT	(CT)7TT(CT)16	60°	—	188
Xgwm52-3D	CTA TGA GGC GGA GGT TGA AG	TGC GCT GCT CTT CCA TTT	(GT)4AT(GT)20	60°	142	128
Xgwm55.1-2B	GCA TCT GGT ACA CTA GCT GCC	TCA TGG ATG CAT CAC ATC CT 3	(TC)3(T)3(CT)17	60°	122	118
Xgwm55.2-2B	GCA TCT GGT ACA CTA GCT GCC	TCA TGG ATG CAT CAC ATC CT 3	(TC)3(T)3(CT)17	60°	161	149
Xgwm55-6D	GCA TCT GGT ACA CTA GCT GCC	TCA TGG ATG CAT CAC ATC CT 3	(TC)3(T)3(CT)17	60°	128	132
Xgwm60-7A	TGT CCT ACA CGG ACC ACG T	GCA TTG ACA GAT GCA CAC G	(CA)30	60°	190	224
Xgwm63-7A	TCG ACC TGA TCG CCC CTA	CGC CCT GGG TGA TGA ATA GT	(CA)17(TA)21	60°	269	271
Xgwm66-4B	CCA AAG ACT GCC ATC TTT CA	CAT GAC TAG CTA GGG TGT GAC A	(CA)30(TA)21	60°	—	218
Xgwm66-5B	CCA AAG ACT GCC ATC TTT CA	CAT GAC TAG CTA GGG TGT GAC A	(CA)30(TA)21	60°	158	137
Xgwm67-5B	ACC ACA CAA ACA AGG TAA GCG	CAA CCC TCT TAA TTT TGT TGG G	(CA)10	60°	94	92
Xgwm68-5B	AGG CCA GAA TCT GGG AAT G	CTC CCT AGA TGG GAG AAG GG	(GA)3(G)3(GA)25	60°	—	166
Xgwm68-7B	AGG CCA GAA TCT GGG AAT G	CTC CCT AGA TGG GAG AAG GG	(GA)3(G)3(GA)25	60°	—	180
Xgwm70-6B	AGT GGC TGG GAG AGT GTC AT	GCC CAT TAC CGA GGA CAC	(GT)7GC(GT)11	60°	197	194
Xgwm71.1-2A	GGC AGA GCA GCG AGA CTC	CAA GTG GAG CAT TAG GTA CAC G	(GT)20	60°	126	124
Xgwm71.2-2A	GGC AGA GCA GCG AGA CTC	CAA GTG GAG CAT TAG GTA CAC G	(GT)20	60°	120	118
Xgwm71-3D	GGC AGA GCA GCG AGA CTC	CAA GTG GAG CAT TAG GTA CAC G	(GT)20	60°	—	101

(continued)

RFLP





Simplified Flowchart of RAC Breeding Program

Triticeae, the wheat tribe

- x = 7
- ca. 22,000 genes
- Conservation of DNA sequences within genes
- Conservation of gene order
- Several diploid species, each with a genome symbol

Diploid wheat relatives

- *T. urartu* A^u
- *T. monococcum* A^m
- *Aegilops* sp. B, G
- *Ae. caudata* C
- *Ae. tauschii* D
- *Agropyron* spp. E, J
- barley H
- *Ae. comosa* M
- *Ae. uniaristata* N
- rye R
- *Ae. speltoides* S
- *Ae. bicornis* S^b
- *Ae. longissima* S^l
- *Ae. mutica* T
- *Ae. umbellulata* U
- *Aegilops* sp. X

Genome evolution in the S group

Formation of polyploids

- A^u pollen met B ovule
- Chromosome doubling required
- 7A + 7A --> 7 II --> meiosis OK
- 7A + 7B --> 14 I --> meiosis fails --> sterile
- "restitution"
- 14AB meets 14AB --> 14 II --> meiosis OK
- *T. turgidum*, a "segmental tetraploid"

Tetraploid wheats

	B, G, S	C	M	N	R	X
A	AB, AG				*AR*	
D		DC	DM	DN	*DR*	
U	US	UC	UM			UX

Hexaploids

	AB	AG	DM	UX
A		AAG		
S			DMS	
D	ABD		DDM	
N				UXN
U			DMU	
R, H	*ABR, ABH*			