Title

Wheat genomics

Journal

Plant Physiology and Biochemistry, 39(3-4)

ISSN

0981-9428

Authors

Lagudah, Evans S
Dubcovsky, Jorge
Powell, Wayne

Publication Date

2001-03-01

DOI

10.1016/s0981-9428(00)01233-x

Peer reviewed

Plant Physiol. Biochem. 39 (2001) 1−10

S098194280001233X/REV

Review

Wheat genomics

Evans Lagudah[a]*, Jorge Dubcovsky[b], Wayne Powell[c§]

[a] CSIRO Plant Industry, GPO Box 1600, Canberra, ACT 2601, Australia

[b] Department of Agronomy, University of California, Davis CA, USA

[c] DuPont Ag Biotechnology, Newark, DE, USA

Received 17 October 2000; accepted 20 November 2000

Abstract – Wheat is the largest produced grain crop world-wide and has been extensively studied for a wide range of agronomic traits located across the genome. Its large chromosomes and the capacity of the polyploid genome to tolerate the addition or loss of chromosomes facilitated a fast progress in early wheat genetics using cytogenetic techniques. However, these same characteristics have limited the progress in genomic studies that have been focused on diploid model species with smaller genomes such as rice and *Arabidopsis*. These model systems are seen as a launching pad for comparative genomic strategies to tackle the challenges associated with the larger genomes of other plants such as wheat. The strengths and limitations of comparative genomic approaches to study the wheat genome are presented in this review. Structural genomic studies in wheat, started by the development of large DNA insert libraries (BAC libraries) for the progenitor A and D genomes using diploid species and for the B genome using tetraploid wheat. Sequencing of the first wheat BAC clones have confirmed cytogenetic results indicating the presence of portions of the wheat genome that contain high density gene regions approximating that of the model species. These high-density gene regions also exhibited higher-than-average rates of genetic to physical distances. Functional genomic approaches, focused on the expressed portion of the wheat genome, have recently led to an exponential growth of expressed sequence tagged (EST) databases from both the public and private sector. Assigning gene function to these ESTs is now one of the major challenges in wheat genomics. Hypothesis testing strategies based on high-resolution EST mapping, candidate gene analysis, gene expression profiling, and proteomics are being used as entry points towards assigning gene function. Improving transformation efficiency as well as selection strategies for high throughput mutagenesis experiments will be two critical areas of research for the ultimate assignment of function to the numerous wheat genes that are being discovered using the new genomic tools. © 2001 Éditions scientifiques et médicales Elsevier SAS

Arabidopsis / colinearity / comparative genomics / functional genomics / rice / wheat

BAC, bacterial artificial chromosome / EST, expressed sequence tag

1. INTRODUCTION

Hexaploid wheat consists of three closely related genomes designated A, B and D which are derived from three progenitor species. Gene redundancy is therefore the norm, with at least a triplicate homoeoallelic set for most of the genes (some genes show deletions for one or more genome). Unlike the smaller genome size of 130–140 Mb in the model experimental plant, *Arabidopsis thaliana* or 400 Mb for the model cereal, rice, wheat has a large genome estimated at 16 000 Mb [1]. In addition to a requirement for smaller genomes in model systems, a significant component of gene discovery strategies relies on the ease of plant transformation. Because wheat lags behind in these two major requirements, genomic studies in wheat are often viewed with considerable challenges. However, the opportunities for the application of genomic studies can be immediate. For example, wheat has an extensive catalogue of over a thousand natural and induced variants that are well characterised for morphological, developmental, biochemical and disease resistant phenotypes [28]. Functional genomic studies when combined with the phenotypic variants at a given locus will have a ready application in wheat breeding.

**Correspondence and reprints:* fax +61 2 62465000.

E-mail address: lagudah@pi.csiro.au (E. Lagudah).

§ Present address: Scottish Crops Research Institute, Dundee, Scotland.

2. COMPARATIVE GENOMICS

Remarkable similarities among eukaryotic genomes provide the basis for using model systems from smaller genomes to examine the structure and organisation of larger genomes such as wheat. In plants, a basic assumption is that all the critical features for gene function associated with growth and development found in the genomics model plant species will be similar in a wide range of plants including crops. Comparative genomic approaches can be exploited either at the level of gene or deduced gene product similarity, or at the level of linear order of genes in the genome (colinearity). The selection between these approaches depends on the level of genetic relatedness between species under investigation. Rice, for example, will be a more appropriate model for wheat where a comparative genomic approach based on conserved micro-synteny is to be used for gene isolation. On the other hand, the vast array of mutants in *Arabidopsis* are key resources that serve as starting points for a gene product-similarity search for analogous wheat traits.

2.1. Gene discovery via sequence similarity

Cloning of the wheat gibberellin insensitive mutants, *Rht-B1* and *Rht-D1*, with dwarfed phenotypes of major impact on grain yield gains in some wheat cultivars, was essentially achieved via homology to similar mutants in *Arabidopsis*. A gibberellin insensitive gene (*GAI*) that encodes for a nuclear transcription factor containing a phosphotyrosine signalling domain first isolated in *Arabidopsis* was used in database searches to identify a rice EST with a high similarity to the N-terminus of *GAI*. Genomic DNA hybridisation of wheat using the rice EST revealed RFLPs that co-segregated with the dwarfing genes *Rht-B1* and *Rht-D1* located on homoeologous groups 4B and 4D chromosomes [33]. Dwarfing mutants that represent alleles of the *Rht-B1* and *Rht-D1* locus were shown to be orthologues of the *GAI* gene from *Arabidopsis*.

Other attempts at isolating genes from *Arabidopsis* homologues have not been so successful. Attempts at isolating wheat or barley photoperiod (*Ppd*) and vernalisation (*Vrn*) response traits based on similarity to the Arabidopsis flowering time gene CONSTANS (*CO*) [34] and FLC [29] have so far proved unsuccessful. However, recent findings from the Rice Genome Program in Japan suggest that a *CO* homologue corresponds to a photoperiod response gene in rice [51].

2.2. Gene discovery via micro-colinearity

Considerable effort has gone into analysing syntenic relationships among grasses [10], and in particular their relationship to the rice genome. During the last decade, blocks of colinear genes were found between wheat and rice and it was thought that information derived from rice genomics would be sufficient to address problems in wheat. Even though information coming from the future complete sequence of the rice genome will be extremely useful for genetic studies in *Triticeae*, its adequacy in tackling many of the critical genetic problems specific to wheat are yet to be ascertained. A few obvious examples where rice information may be of limited value are breadmaking and pasta quality, winter hardiness, vernalisation requirement, etc. An additional limitation in the use of rice information for positional cloning efforts in the Triticeae is that numerous cDNA clones have been shown to hybridise to non-colinear regions in wheat and rice. A possible explanation for this lack of colinearity is the higher level of intra- and inter-chromosomal gene duplication observed in wheat maps compared to rice maps [13]. An additional observation is that many multigene families including resistance genes [30, 45], storage proteins and ribosomal genes, are in non-homoeologous loci even in species within the *Triticeae* [11–13]. The substantial evolutionary divergence between rice and *Triticeae* species increases the possibility of finding independent duplications or non-homologous multigene families, and therefore, the possibility that a targeted gene is not present in wheat and rice homoeologous regions.

Despite these limitations some attempts have been made to use rice as a stepping stone to clone wheat genes. A 300-kb sequenced region of rice chromosome 9 reveals a high level of synteny to wheat chromosome 5BL containing the *Ph1* gene that controls strict homologous chromosome pairing [36]. Although colinearity was observed for several markers, segments of some of the rice markers occurred outside the target region. Cross hybridisation of markers from rice BACs to the target region in wheat in combination with new *Ph1* deletion mutants was used to further narrow the *Ph1* locus to an equivalent region of ~150 kb in rice. Ten predicted genes were identified in the 150-kb region, some of which included transcription factors, receptor kinases, metabolic enzymes and anonymous ESTs; however, it remains to be demonstrated whether any of the predicted genes correspond to a candidate gene for *Ph1* [18]. In attempts at positional cloning of the barley stem rust resistance gene, *Rpg1*, located on the distal end of

barley chromosome 7HS which corresponds to homoeologous group 7S in wheat, a high density map was constructed around the *Rpg1* region aided by additional markers from a colinear region from rice. A 70-kb rice BAC clone carrying markers that flanked the predicted *Rpg1* homologue was completely sequenced [19]. None of the gene sequences from the rice BAC showed features expected of disease resistance genes. In contrast, homologues of the maize rust resistance gene *Rp1* as well as other disease resistance gene homologues have been shown to be tightly linked to *Rpg1* [2, 9]. The absence of these disease resistance gene homologues in the corresponding region in rice may be attributed to the rapid re-organisation of these genes in plant genomes [23].

Lack of colinearity in resistance genes has also been observed among closely related wheat species. A homoeologous segment from the tetraploid wheat relative, *Aegilops ventricosa* (DN gemome), carrying stem, stripe and leaf rust resistance genes as well as a family of disease resistance gene homologues (Vrga1) was introgressed onto chromosome 2A [38, 39]. Vrgal family members were completely absent from the D genome of the *Ae. ventricosa* donor. Subsequent analysis of a random set of diploid D genome donor lines of wheat revealed that null haplotypes (or complete absence) of the Vrga1 family members were frequent among the accessions (*figure 1*). Such diverse organisation of predicted disease resistance genes and their frequent occurrence at non-syntenic loci even within the wheat species suggests confounding effects will arise when comparative genomic approaches are used. Sole reliance on comparative genomics approaches based on colinearity between rice and the wheat genomes in cloning disease resistance genes, and possibly other multigene families, may thus prove to be elusive.

2.3. Genetic filtration and gene discovery

Other alternative strategies aimed at filtering large genomes in order to develop gene enriched genomic libraries have been applied to maize. Between 60–80 % of the maize genome consists of repetitive elements [6], which are similar to wheat. The retrotransposons and other repeated sequences are highly methylated, and can be excluded from genomic libraries ('filtered') by exploiting the use of methylation restrictive *Escherichia coli* host strains. A four-fold increase in the proportion of known coding gene sequences from GenBank database was demonstrated with maize filtered over the unfiltered library [35]. Genetic filtering has therefore been suggested as a way of achieving comprehensive gene discovery in wheat through genomic sequencing when combined with the imminent complete genome references from Arabidopsis and rice. It remains to be shown what proportion of exons or inter-repeat and gene regions may be excluded in genetically filtered genomic libraries.

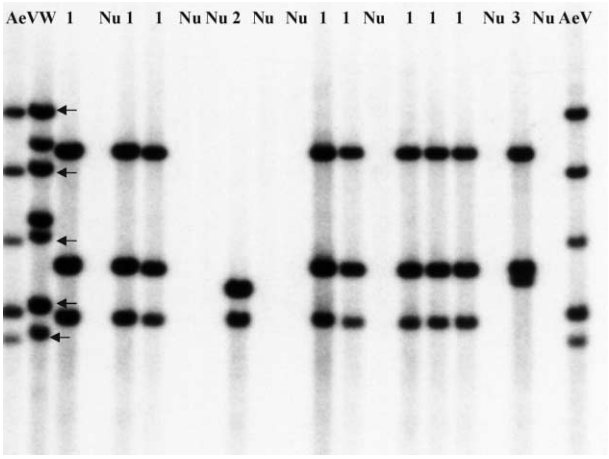

Figure 1. RFLP haplotypes of a disease resistance gene analogue (Vrga1) linked to rust resistance genes in the distal region on homoeologous group 2 chromosomes. AeV, *Aegilops ventricosa*; W, bread wheat cultivar carrying Vrga1 members transferred from *Ae. ventricosa*; 1, 2, 3 and Nu represent four different haplotypes present in *Ae. tauschii* –the diploid D genome donor. The arrowed fragments are Vrga1 gene members derived only from the N genome of the tetraploid (DN genome) donor species *Ae. ventricosa*. The D genome of the *Ae. ventricosa* line lacks Vrga1 gene members. About 40 % of the random set of diploid D genome species carry null (Nu) phenotypes; an indicator of the rapid re-organisation of the Vrga1 gene region.

3. STRUCTURAL GENOMICS

3.1. Physical maps based on deletion lines

Wheat is an hexaploid species and therefore tolerates chromosome deletions and aneuploidy better than diploid organisms. Genetic stocks missing complete chromosomes or chromosome arms were developed by E. Sears [40] and have greatly facilitated gene synteny mapping for almost 50 years. Advances in C-banding techniques and the discovery of a genetic system that generates unlimited numbers of terminal deletions opened an era of sub-arm localisation of genes to specific chromosome segments [15]. RFLP probe hybridised with DNA from these lines show restriction fragments present in all the lines with deletions distal to the locus but absent in all lines with deletions proximal to the corresponding loci (*figure 2*)

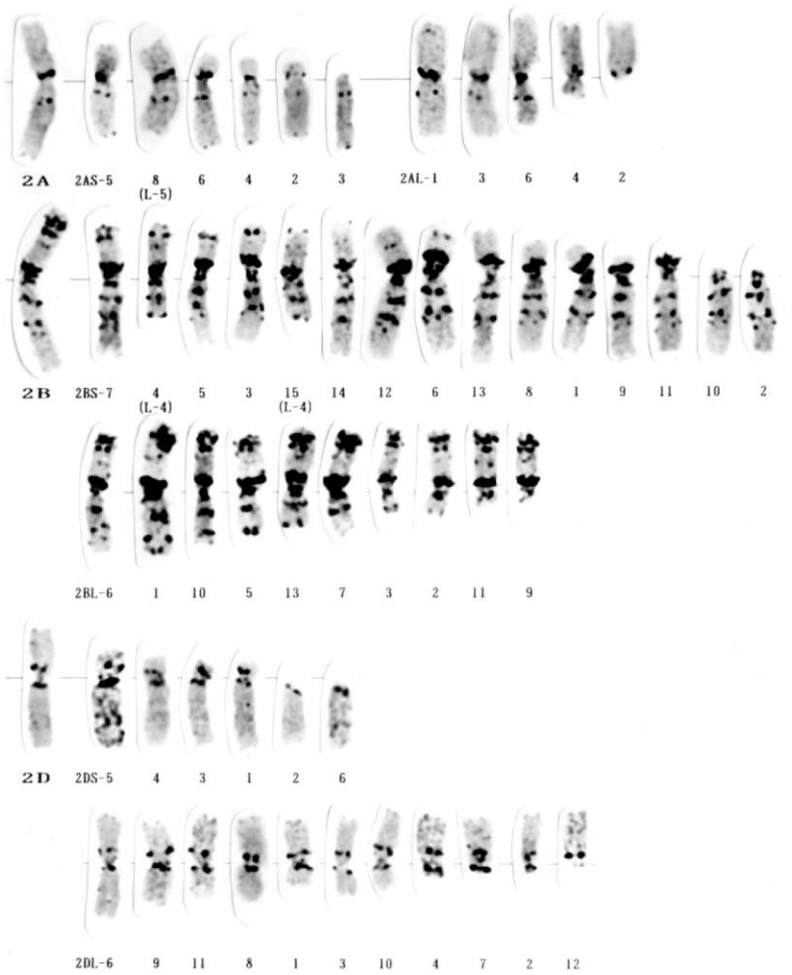

Figure 2. A deletion series of homoeologous group 2 chromosomes of the bread wheat variety, Chinese Spring (kindly provided by T.R. Endo) to facilitate gene location [15, 17, 48]. Rapid localisation of genes or DNA fragments into a chromosomal bin is achieved by the presence or absence of an RFLP when the corresponding intact or loss of chromosome segment occurs.

Most fragments detected in genomic DNA hybridisation of deletion lines are efficiently mapped in a single hybridisation without requiring a preliminary screening for polymorphisms. The simplicity of this process has facilitated the assignment of more than 300 RFLP markers to 436 deletions distributed along the 21 chromosomes of wheat [17, 49]. A consortium of thirteen USA-laboratories has been recently funded to assign 10 000 unique ESTs to a subset of these chromosome bins (http://wheat.pw.usda.gov/NSF/home.html).

The combination of deletion lines from the three homoeologous chromosomes divides each consensus chromosome into approximate 62 different 'physical bins' with an average size of 40 Mb. Assuming a uniform distribution of recombination, these bins would represent segments of 10 cM in the genetic maps (average ratio of physical to genetic distances 4 $Mb \cdot cM^{-1}$). However, comparisons of physical maps with linkage maps using common sets of RFLP markers have confirmed that centromeric regions have lower levels of recombination than distal regions of the wheat chromosomes.

Analysis of the distribution of cDNA markers along the different deletion lines has also shown a non-uniform distribution. Absence or reduced number of cDNA markers in the centromeric deletions parallels the absence of recombination in these chromosome regions [16, 17]. Some small deletions showed extremely high number of cDNA markers suggesting the presence of 'gene islands' within the wheat genome.

3.2. Large insert libraries

Chromosome segments defined by adjacent deletions provide a first level approximation of physical maps of wheat chromosomes. A higher level of reso-

lution, appropriate for positional cloning experiments, was recently achieved by the use of bacterial artificial chromosome clones (BAC) [41]. BAC clones can maintain an exogenous DNA fragment up to 300 kb in size in *E. coli* cells and are relatively easy to manipulate. BAC libraries have a long-term value because they can be stored as individual clones and can permanently serve as a genomic research resource. Three different BAC libraries have been constructed from diploid wheat *Triticum tauschii* (syn. *Aegilops tauschii*; D genome), *T. monococcum* (A^m genome) and the bread wheat cultivar, Chinese Spring. Additional libraries are under construction for pasta (AB, genomes) and other bread wheats (ABD genomes).

The *T. tauschii* large DNA insert library consists of 144 000 clones with an average insert size of 119 kb, and was estimated to be equivalent to 3.7 haploid genomes [31]. This library has been blotted into 3 474 clones per high density filters and is publicly available. The library was constructed using a binary cosmid and one of the clones (80-kb insert) tested in rice transformation experiments was shown to be stably maintained. The *T. monococcum* BAC library consists of 276 480 clones with an average insert size of 115 kb. Excluding the 1.33 % of empty clones and 0.14 % of clones with chloroplast DNA, the coverage of this library is 5.6 genome equivalents [24]. With this genome coverage, the probability of having any DNA sequence represented in this library is higher than 99.6 %. Clones from this library are sorted in 720 384 well plates and blotted onto fifteen high-density filters (18 432 double spotted clones per filter). High-density filters and clones are also publicly available (http://agronomy.ucdavis.edu/Dubcovsky/BAC-library/BAC_Home.htm).

A tetraploid BAC library is being constructed at the Wheat Genomic Center (University of California, Davis & USDA, Albany California). Half of a million BAC clones (average insert size 125 kb) were produced, picked and arrayed in 384 well plates. The status of this library can be found at http://agronomy.ucdavis.edu/Dubcovsky/BAC-library/BAC_Langdon.htm. The hexaploid wheat library from Chinese Spring was constructed in a newly developed transformation-competent artificial chromosome (TAC) vector, pYLTAC17 [25]. Unlike the other libraries, clones from the Chinese Spring hexaploid wheat library are stored as pools of 300–600 colonies per micro-titre well rather than as single colonies. Other research programmes currently working on the construction of BAC libraries from hexaploid wheat include the John Innes Centre (UK) also on the wheat variety 'Chinese Spring'; Agriculture Canada on the wheat variety 'Glenlea' and the CSIRO (Australia) on the variety 'Hartog'.

3.3. Global contigs

An ideal physical map of a chromosome consists of a set of minimum overlapping BAC clones covering the complete chromosome and including numerous connections to linkage maps. After a trait is linked to the genetic map, the flanking markers can be used to determine the position of the region in the physical map. BACs from this region can then be used to saturate the targeted region with molecular markers, to explore the presence of candidate genes, and, if necessary for chromosome walking. The construction of BAC contigs for the wheat chromosomes has the potential to transform the most difficult part of a positional cloning project into a trivial bioinformatic exercise.

The large size of the wheat genome and the abundance of large repetitive elements create additional levels of complexity to the already difficult task of assembling complete chromosome contigs. The magnitude of this task becomes clear when one realises that a single chromosome of wheat contains more DNA than the complete haploid genome of rice. Fortunately, new BAC fingerprinting techniques, based on accurate sizing of DNA fragments (0.15 bp) on automated capillary 3100 and 3700 ABI platforms, provide a better resolution and automation possibilities than the currently used agarose gel-base fingerprinting techniques [27]. This technology is currently being used at DuPont® to assemble a contig of the maize genome and has been selected to assemble BAC contigs for diploid wheat by an NSF-USA project (J. Dvorak PI). Diploid *T. tauschii* was selected for this project because of its relatively smaller genome size (4 700 Mb per haploid genome) compared with the other wheat diploid species involved in the origin of polyploid bread wheat.

3.4. Local contigs

Even though global contigs for the complete wheat chromosomes are still not available, local contigs constructed in different chromosomes by different laboratories provide preliminary estimates of the relationship between physical and genetic distances in these regions. Since all these BACs were selected by cDNA clones, these estimates should be considered representative of regions including at least one gene.

The regions studied so far, include the distal region of chromosomes 1AS = 420 kb·cM^{-1} [20], 1DS =

Table I. Base pair composition of *Copia*-like retroelements and wheat BAC clones

	A	T	C	G	GC
1A^{m}S BAC	27.0	26.8	22.8	23.5	46.3
5A^{m}L BAC	25.2	27.3	23.3	24.1	47.5
Wis-2-A1 from 1A^{m}S BAC	28.2	27.3	21.0	23.5	44.5
Barley BARE1 (GI397873)	25.7	28.1	21.4	24.8	46.2
Nebulised genomic library	27.2	27.1	22.7	23.0	45.7

20–270 kb·cM^{-1} [44], 5A^{m}S = 260 kb·cM^{-1} [47]; the proximal region of chromosome 1AS (J. Dubcovsky and W. Powell, unpubl. results, 250 kb·cM^{-1}); and the central region of 5A^{m}L (J. Dubcovsky and J. Bennetzen, unpubl., 1 970 kb·cM^{-1}). These estimates of physical distances per unit of recombination are 2 to 150 times lower than the genome-wide estimate of 3 000 kb·cM^{-1} for wheat [5].

3.5. Gene density

The relatively high frequency of recombination observed between markers located within these wheat BACs parallels their relatively high gene density. If the total number of genes in diploid wheat were similar to the estimated 21 000 genes in *A. thaliana* [7] and the genes were randomly distributed along the genome, the expected average distance between genes would be approximately 250 kb. However, gene densities observed in this set of wheat BACs were 7- to 40-fold higher than expected based on a random gene distribution ([20, 47]; J. Dubcovsky, W. Powell and E.S. Lagudah, unpubl.). These estimates agree with previous observations based on wheat deletion lines [17] and from the sequencing of a 60-kb region in barley [32] and all suggest that genes are organised in gene-rich regions or gene islands [3] within the large genome of the *Triticeae* species.

3.6. Sequence analysis of selected and random genomic regions

Two BACs selected with clones corresponding to genes located on chromosomes 1A^{m}S (J. Dubcovsky and W. Powell, unpubl.) and 5A^{m}L (W. Ramakrishna, P. SanMiguel, J. Dubcovsky and J. Bennetzen, unpubl.) of *Triticum monococcum* have been recently sequenced. These first sequences showed two to four genes per BAC interspersed among repetitive regions. The repetitive elements found in these BACs comprise mainly *Copia*-like retroelements and fragments of other repetitive sequences. *Copia*-like retroelements account for approximately one-fifth of the sequence of each of these two BACs. Not surprisingly, the base pair composition of these two BACs matches closely the base pair composition of *Copia*-like retroelements (*table I*). This base pair composition is also similar to the average base pair composition of a random sample of 1.02 Mb of sequencing from a nebulised genomic library of *T. aestivum* generated at Dupont® (*figure 3*; *table I*). This result suggests that these two gene-rich BACs have a base pair composition similar to the genome average.

The proportion of *Copia*-like elements in the 1.02-Mb random sample of *T. aestivum* genomic DNA (13 %; *figure 3*) is slightly smaller than the proportion found in 1A^{m}S BAC clone (22.5 %) and the 5A^{m}L BAC clone (17.5 %) suggesting the possibility of an enrichment of *Copia*-like retroelements in these particular gene regions.

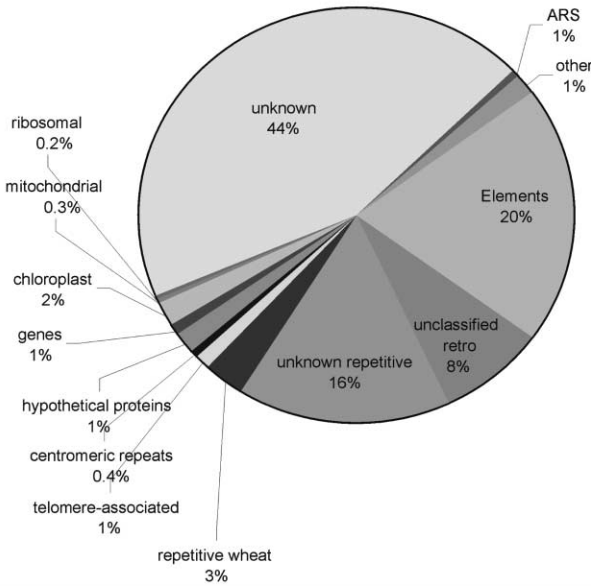

Figure 3. A summary of the distribution of different sequences found from random sequencing of a nebulised bread wheat genomic library. A total of 1 016-kb sequences were obtained (W. Powell, K. Henderson and P. Wolters, unpubl.). Out of the 20 % of repeat elements, 13 % were *Copia*-like elements.

4. FUNCTIONAL GENOMICS

4.1. ESTs

The basic strategy for EST (expressed sequence tags) production and use was formulated by Craig Venter's group (TIGR) in 1991 and is a rapid efficient method for sampling a genome for active gene sequences. Typically, anonymous cDNAs are used to determine short DNA sequences (300–400 bp) in a single sequencing reaction. These sequences are then used as tags to search existing databases to determine if a specific gene (or gene motif) has been found in the same or other organisms and if its function has been determined. By far most of the advances towards the initial steps of functional genomics in wheat have been the generation of large numbers of ESTs from different tissues. The recent sharp increases in the number of ESTs from wheat (from > 100 to < 29 000) available in GenBank stems from a co-ordinated effort of public institutions, under the auspices of the International Triticeae EST Consortium (http://wheat.pw.usda.gov/genome/index.html). A much larger number of ESTs outside GenBank are held in several national institutions as well as in the private sector, which includes > 125 000 in Du Pont's database. Most of the available ESTs are from tissues that include the developing grain, leaf, root, stress (cold, heat, disease) induced and un-induced seedlings. ESTs from the wheat grain appear to be over-represented; a reflection of their importance in gene discovery for wheat products.

4.2. Approximating gene function

Determining gene function, undoubtedly remains the biggest challenge in wheat genomics. A number of hypothesis generation methods are being used towards deducing gene function. As a first step, homology of wheat ESTs or genomic clones with presumptive orthologous genes from other plants is established. If the function of a gene in other plant species is known a similar function is suggested for the wheat EST. A second step is the linkage of the EST with cognate phenotypes in mapping families. High-density maps can then be used to increase the strength of the linkage hypothesis. However, the non-random distribution of suppressed recombination and hot spots of recombination in the wheat genome, as well as the presence of multigene families at a single locus in wheat, pose limitations to the candidate gene analysis through linkage.

Another way to approximate gene function for thousands of genes simultaneously is the assembly of large number of wheat ESTs on solid media (micro-arrays) to facilitate global gene expression studies in different tissues, genotypes and environments. For example, micro-arrays of grain ESTs will facilitate the integration of our current knowledge of the expression patterns of genes affecting starch, gluten, and lipids synthesis under different environmental conditions and their combined effects on product quality as is often found during grain development and maturation in the field.

The use of high-resolution 2-dimensional protein electrophoresis and improvements in heterologous systems for expressing isolated genes from the wheat grain will provide a proteomics approach to complement functional genomics studies of the wheat grain. Modifications to conventional micro-arrays as well as a proteomic approach are currently being used in an attempt to identify novel genes that may be implicated in the differential initiation and synthesis of the wheat starch A- and B-granules [4].

4.3. Validation of gene function

The ultimate proof of gene function is reliant on either gene complementation or the development of mutants where alterations to the plant's phenotype can be directly linked to specific changes in the wild type gene. Most of the elegant systems developed for gene discovery such as enhancer, promoter and gene traps in model plants are dependent on highly efficient transformation systems. Herein lies one of the major obstacles to the validation phase of functional genomics in wheat. Current wheat transformation efficiency ranges from mean values between 0.1–5 %, and are too low to develop a transformation-based gene tagging system. Some attempts have been made at transferring the maize *Ac* (Activator) and *Ds* (Dissociation) transposable elements into wheat callus lines. Activation and excision of the maize *Ac* and *Ds* elements was demonstrated in the transgenic wheat calli, and similar to observations in other plants, *Ac* transposase transcripts at high levels resulted in a decreased excision frequency [46]. Thus in principle, gene discovery via heterologous transposon tagging is possible with improvements in wheat transformation and regeneration.

Gene redundancy due to the polyploid nature of the wheat genome is a potential limitation to insertional mutagenesis, complementation studies or the rapid development of chemically induced fine scale mutants. Recent improvements in the development of efficient gene silencing constructs against viruses or endogenous genes based on intron-spliced doubled stranded RNA [42] offer opportunities in overcoming gene redundancy problems in functional genomics. A rela-

tively short stretch of nucleotides (120–700 bp) showing high homology among closely related gene family members may be all that is required to develop an intron-spliced hairpin RNA capable of silencing a gene family. This approach has been successfully applied in developing immunity to barley yellow dwarf virus (BYDV) in barley and is currently being tested in wheat ([48]; L. Rooke, D. Abbott, M.-B. Wang and P.M. Waterhouse, unpubl.).

5. APPLICATIONS: GENOMICS FROM LAB TO FIELD

5.1. Molecular markers

5.1.1. Quality

An immediate application of gene discovery programmes in wheat is the development of perfect markers for genes affecting grain quality. One example was the development of molecular markers for deletion mutants with loss of granule bound starch synthase (GBSS) genes that are now the basis for precision breeding of commercially produced speciality noodles [52]. Another example was the development of markers for starch granule protein-1 (SGP-1) null phenotypes that possess high amylose starch in wheat [50]. Finally, the characterisation of BAC clones spanning the genes responsible for the 'softness' of the grain from the A and D genomes facilitated the development of PCR markers for this trait ([47]; K. Turnbull and S. Rahman, unpubl.).

Availability of molecular markers for the main genes affecting quality has opened new avenues for commercial developments of different wheat products. Varieties with altered starch properties can be selected using markers for the SGP-1 and GBSS mutants. The hardness of the grain can be altered by the introgression of wild 'softness' alleles or by the selection of lines with multiple deletions in the different 'softness' genes (G. Tranquilli and J. Dubcovsky, unpubl.). The strength of the gluten can also be modified using simultaneously molecular markers for high and low molecular weight glutenins [26].

5.1.2. Disease resistance

Another area where wheat genomics can generate new knowledge with immediate applications is the area of disease resistance genes. Gene sequences encoding for nucleotide binding site and leucine-rich repeat (NBS-LRR) sequences are by far the most abundant class of disease resistance genes cloned in plants [14]. An estimate of 200 NBS-LRR sequences is present in the *Arabidopsis* genome [30], and in wheat, homologues of these genes are present on all chromosomes [45]. Because a wide range of disease resistance phenotypes are present across the entire wheat genome, resistance gene homologues identified either through linkage analysis or functional genomics will lead to an increasing number of perfect markers for resistance breeding. An NBS-LRR sequence at the *Yr10* locus on chromosome 1B has been demonstrated by complementation to confer stripe rust resistance in wheat (A. Laroche, pers. comm.). Homologues of the *Yr10* gene present at a syntenic position on wheat chromosome 1D have also been shown to co-segregate with a leaf rust resistance gene, *Lr21* [43], and thus provide an entry point for isolating the *Lr21* gene. Furthermore NBS-LRR sequences at the *Cre3* nematode resistance locus are also present at non-syntenic loci which codes for other cyst nematode resistance genes, *Cre1*, *Cre5* and *Cre6* [21, 22]. Sequences from the *Cre3* locus are currently employed as the basis for developing cereal cyst nematode resistance lines in the CIMMYT and other national wheat breeding programmes.

Finally, genomic studies in wheat are generating numerous simple sequence repeats (SSR) markers that will complement those specifically generated by SSR projects [37]. Recent analyses of SSR repeats from wheat EST libraries are beginning to uncover a wide range of microsatellite variants (W. Powell, K. Henderson and P. Wolters unpubl.). One advantage of EST SSRs is that they serve as functional markers that directly assay the deduced gene from the EST, and that some of them detect homeoalleles in more than one wheat genome. An additional source of SSR markers is the sequence of BAC clones. For example, sequence from the *Triticum monococcum* $1A^m$ BAC that includes the *Triticin* gene was used to develop an SSR marker completely linked to this gene.

5.2. Transgenic approaches

Besides the development of perfect markers, the discovery of new genes will allow the engineering of wheat using transgenic approaches. The viability of this technology in wheat has been demonstrated by the introduction of transgenic storage proteins in wheat cultivars [8]. Variable contents of high molecular weight (HMW) glutenin subunits exceeding those found in natural variants have been identified through either over-expression or silencing of endogenous HMW glutenin genes in transgenic wheat. These new

variants have impacted on dough functional properties, and may lead to new blending schemes that may allow for a more flexible use of flour samples from a wide range of wheat varieties. Cytogenetic manipulation has a long tradition in wheat for transferring disease resistance genes from species within the Triticeae. Some of these gene transfers are accompanied by large segments of the alien chromatin with a resultant linkage drag of undesirable phenotypes. Some successes have been achieved with transferring single resistance genes across diverse taxa in the Solanaceae. It is expected that as additional resistance genes are isolated from other genera within the Triticeae and other grasses, their expression and modification as transgenes in wheat will be evaluated for transformation breeding.

REFERENCES

[1] Arumuganathan K., Earle E.D., Nuclear DNA content of some important plant species, Plant Mol. Biol. Rep. 9 (1991) 208–218.

[2] Ayliffe M.A., Collins N.C., Ellis J.G., Pryor A., The maize *rp1* rust resistance gene identifies homologues in barley that have been subjected to diversifying selection, Theor. Appl. Genet. 100 (2000) 1144–1154.

[3] Barakat A., Carels N., Bernardi G., The distribution of genes in the genomes of Gramineae, Proc. Natl. Acad. Sci. USA 94 (1997) 6857–6861.

[4] Beckles D.M., Broglie K.E., Investigating starch biosynthesis in wheat, International Triticeae Mapping Initiative: Tenth International Public Workshop, 2000, p. 5.

[5] Bennett M.D., Leitch I.J., Nuclear DNA amounts in Angiosperms, Ann. Bot. 76 (1995) 113–176.

[6] Bennetzen J.L., Comparative sequence analysis of plant nuclear genomes: microcolinearity and its many exceptions, Plant Cell 12 (2000) 1021–1029.

[7] Bevan M., Bancroft I., Bent E., Love K., Goodman H., Dean C., et al., Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*, Nature 391 (1998) 485–488.

[8] Blechl A.E., Le H.Q., Anderson O.D., Engineering changes in wheat flour by genetic transformation, J. Plant Physiol. 152 (1998) 703–707.

[9] Collins N., Drake J., Ayliffe M., Qing S., Ellis J., Hulbert S., Pryor T., Molecular characterisation of the maize *Rp1D* rust resistance haplotype and its mutants, Plant Cell 11 (1999) 1–13.

[10] Devos K.M., Gale M.J., Genome relationships: The grass model in current research, Plant Cell 12 (2000) 637–646.

[11] Dubcovsky J., Dvorak J., Ribosomal RNA loci: Nomads in the Triticeae genomes, Genetics 140 (1995) 1367–1377.

[12] Dubcovsky J., Echaide M., Giancola S., Rousset M., Luo M.C., Joppa L.R., Dvorak J., Seed-storage-protein loci in RFLP maps of diploid, tetraploid, and hexaploid wheat, Theor. Appl. Genet. 7 (1997) 1169–1180.

[13] Dubcovsky J., Luo M.C., Zhong G.Y., Bransteitter R., Desai A., Kilian A., Kleinhofs A., Dvorak J., Genetic map of diploid wheat, *Triticum monococcum* L., and its comparison with maps of *Hordeum vulgare* L, Genetics 143 (1996) 983–999.

[14] Ellis J.G., Dodds P., Pryor T., The generation of plant disease resistance specificities, Trends Plant Sci. 5 (2000) 373–379.

[15] Endo T.R., Gill B.S., The deletion stocks of common wheat, J. Hered. 87 (1996) 295–307.

[16] Gill K.S., Gill B.S., Endo T.R., A chromosome region-specific mapping strategy reveals gene-rich telomeric ends in wheat, Chromosoma 102 (1993) 374–381.

[17] Gill K.S., Gill B.S., Endo T.R., Boyko E.V., Identification and high-density mapping of gene-rich regions in chromosome group 5 of wheat, Genetics 143 (1996) 1001–1012.

[18] Griffith S., Dalgliesh C., Foote T.N., Henderson K., Powell W., Moore G., The molecular dissection of the *Ph1* locus, International Triticeae Mapping Initiative: Tenth International Public Workshop, 2000, p. 6.

[19] Han F., Kilian A., Chen J.P., Kudrna D., Steffenson B., Yamamoto K., Matsumoto T., Sasaki T., Kleinhofs A., Sequence analysis of a rice BAC covering the syntenous barley *Rpg1* region, Genome 42 (1999) 1071–1076.

[20] Keller B., Feuillet C., Colinearity and gene density in grass genomes, Trends Plant Sci. 5 (2000) 246–251.

[21] Lagudah E.S., Moullet O., Appels R., Map based cloning of a gene sequence encoding a nucleotide binding domain and leucine rich region at the *Cre3* nematode resistance locus of wheat, Genome 40 (1997) 659–665.

[22] Lagudah E.S., Moullet O., Ogbonnaya F., Seah S., Eastwood R., Appels R., Jahier J., Lopez-Brana I., Delibes A., Cloning of disease resistance gene sequences at loci conferring cyst nematode resistance genes in wheat, in: Slinkard A.E. (Ed.), Proc 9th Int. Wheat Genet. Symp., Saskatoon, Saskachetwan, Canada, 1998, pp. 184–186.

[23] Leister D., Kurth J., Laurie D.A., Yano M., Sasaki T., Devos K., Graner A., Schulze-Lefert P., Rapid reorganization of resistance gene homologues in cereal genomes, Proc. Natl. Acad. Sci. USA 95 (1998) 370–375.

[24] Lijavetzky D., Muzzi G., Wicker T., Keller B., Wing R., Dubcovsky J., Construction and characterization of a bacterial artificial chromosome (BAC) library for the A genome of wheat, Genome 42 (1999) 1176–1182.

[25] Liu Y.G., Nagaki K., Fujita M., Kawaura K., Uozumi M., Ogihara Y., Development of an efficient maintenance and screening system for large-insert genomic DNA libraries of hexaploid wheat in a transformation-competent artificial chromosome (TAC) vector, Plant J. 23 (2000) 687–695.

[26] Manifesto M.M., Feingold S., Hopp H.E., Schlat-

ter A.R., Dubcovsky J., Molecular markers associated with differences in breadmaking quality in a cross between bread wheat cultivars with the same HMr glutenins, J. Cereal Sci. 27 (1998) 217–227.
[27] Marra M.A., Kucaba T.A., Dietrich N.L., Green E.D., Brownstein B., Wilson R.K., McDonald K.M., Hillier L.W., McPherson J.D., Waterston R.H., High throughput fingerprint analysis of large-insert clones, Genome Res 7 (1997) 1072–1084.
[28] McIntosh R.A., Devos K.M., Dubcovsky J., Rogers W.J., Catalogue of gene symbols for wheat (2000)http://wheat.pw.usda.gov/ggpages/wgc/2000upd.html.
[29] Michaels S.D., Amasino R.M., Flowering locus C encodes a novel MADS domain protein that acts as a repressor of flowering, Plant Cell 11 (1999) 949–956.
[30] Michelmore R., Genomic approaches to plant disease resistance, Curr. Opin. Plant Biol. 3 (2000) 125–131.
[31] Moullet O., Zhang H.B., Lagudah E.S., Construction and characterization of a large DNA insert library from the D genome of wheat, Theor. Appl. Genet. 99 (1999) 305–313.
[32] Panstruga R., Buschges R., Schulze-Lefert P., A contiguous 60 kb genomic stretch from barley reveals molecular evidence for gene islands in a monocot genome, Nucleic Acids Res. 26 (1998) 1056–1062.
[33] Peng J.R., Richards D.E., Hartley N.M., Murphy G.P., Devos K.M., Flintham J.E., Beales J., Fish L.J., Worland A.J., Pelica F., Sudhakar D., Christou P., Snape J.W., Gale M.D., Harberd N.P., 'Green revolution' genes encode mutant gibberellin response modulators, Nature 400 (1999) 256–261.
[34] Putteril J., Robson F., Lee K., Simon R., Coupland G., The CONSTANS gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors, Cell 80 (1995) 847–857.
[35] Rabinowicz P.D., Schutz K., Dedhia N., Yordan C., Parnell L.D., Stein L., McCombie W.R., Martienssen R.A., Differential methylation of genes and retrotransposons facilitates shotgun sequencing of the maize genome, Nat. Genet. 23 (1999) 305–308.
[36] Roberts M.A., Reader S.M., Dalgliesh C., Miller T.E., Foote T.N., Fish L.J., Snape J.W., Moore G., Induction and characterization of Ph1 wheat mutants, Genetics 153 (1999) 1909–1918.
[37] Roder M.S., Korzun V., Wendehake K., Plaschke J., Tixier M.H., Leroy P., Ganal M.W., A microsatellite map of wheat, Genetics 149 (1998) 2007–2023.
[38] Seah S., Bariana H., Jahier J., Sivasithamparam K., Lagudah E.S., The introgressed segment carrying rust resistance genes *Yr17*, *Lr37* and *Sr38* in wheat can be assayed by a cloned disease resistance gene-like sequence, Theor. Appl. Genet. 101 (2000).
[39] Seah S., Spielmeyer W., Jahier J., Sivasithamparam K., Lagudah E.S., Resistance gene analogs within an introgressed chromosomal segment derived from *Triticum ventricosum* that confers resistance to nematode and rust pathogens in wheat, Mol. Plant Microbe-Interact. 13 (2000) 334–341.
[40] Sears E.R., The aneuploids of common wheat, Res. Bull. Univ. Missouri Agric. Exp. Stn. 572 (1954) 1–59.
[41] Shizuya H., Birren B., Kim U.J., Mancino V., Slepak T., Tachiiri Y., Simon M., Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector, Proc. Natl. Acad. Sci. USA 89 (1992) 8794–8797.
[42] Smith N.A., Singh S.P., Wang M.B., Stoutjesdijk P.A., Green A.G., Waterhouse P.M., Gene expression - Total silencing by intron-spliced hairpin RNAs, Nature 407 (2000) 319–320.
[43] Spielmeyer W., Huang L., Bariana H., Laroche A., Gill B.S., Lagudah E.S., NBS-LRR sequence family is associated with leaf and stripe rust resistance on the end of homoeologous chromosome group 1S of wheat, Theor. Appl. Genet. 101 (2000) 1139–1144.
[44] Spielmeyer W., Moullet O., Laroche A., Lagudah E.S., Highly recombinogenic regions at seed storage protein loci on chromosome 1DS of *Aegilops tauschii*, the D-genome donor of wheat, Genetics 155 (2000) 361–367.
[45] Spielmeyer W., Robertson M., Collins N., Leister D., Schulze-Lefert P., Seah S., Moullet O., Lagudah E.S., A superfamily of disease resistance gene analogs is located on all homoeologous chromosome groups of wheat (*Triticum aestivum*), Genome 41 (1998) 782–788.
[46] Takumi S., Murai K., Mori K., Nakamura C., Variations in the maize *Ac* transposase transcript level and the *Ds* excision frequency in transgenic wheat callus lines, Genome 42 (1999) 1234–1241.
[47] Tranquilli G., Lijavetzky D., Muzzi G., Dubcovsky J., Genetic and physical characterization of grain texture-related loci in diploid wheat, Mol. Gen. Genet. 262 (1999) 846–850.
[48] Wang M.B., Abbott D.C., Waterhouse P.M., A single copy of a virus-derived transgene encoding hairpin RNA gives immunity to barley yellow dwarf virus, Mol. Plant Pathol. 1 (2000).
[49] Werner J.E., Endo T.R., Gill B.S., Towards a cytogenetically based physical map of the wheat genome, Proc. Natl. Acad. Sci. USA 89 (1992) 11307–11311.
[50] Yamamori M., Fujita S., Hayakawa K., Matsuki J., Yasui T., Genetic elimination of a starch granule protein, SGP-1, of wheat generates an altered starch with apparent high amylose, Theor. Appl. Genet. 101 (2000) 21–29.
[51] Yano M., Katayose Y., Ashikari M., Yamanouchi U., Monna L., Fuse T., Baba T., Yamamoto K., Umehara Y., Nagamura Y., Takuji Sasaki T., Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering-time gene CONSTANS, Plant Cell 12 (2000).
[52] Zhao X.C., Batey I.L., Sharp P.J., Crosbie G., Barclay I., Wison R., Morell M.K., Appels R., A single genetic locus associated with starch granule properties and noodle quality in wheat, J. Cereal Sci. 27 (1998) 7–13.