

To clone or not to clone plant QTLs: present and future challenges

Silvio Salvi and Roberto Tuberosa

Department of Agroenvironmental Science and Technology, University of Bologna, Viale Fanin, 44, 40127 Bologna, Italy

Recent technical advancements and refinement of analytical methods have enabled the loci (quantitative trait loci, QTLs) responsible for the genetic control of quantitative traits to be dissected molecularly. To date, most plant QTLs have been cloned using a positional cloning approach following identification in experimental crosses. In some cases, an association between sequence variation at a candidate gene and a phenotype has been established by analysing existing genetic accessions. These strategies can be refined using appropriate genetic materials and the latest developments in genomics platforms. We foresee that although QTL analysis and cloning addressing naturally occurring genetic variation should shed light on mechanisms of plant adaptation, a greater emphasis on approaches relying on mutagenesis and candidate gene validation is likely to accelerate the pace of discovering the genes underlying QTLs.

From polygenes to QTL cloning

Classical quantitative geneticists defined 'polygenes' [1] as the many hypothetical genes, with an equally small effect, involved in determining a quantitative trait (i.e. a trait influenced by both multiple genetic and environmental factors). Polygenes have been integrated into most quantitative genetics models; in many cases, these models successfully describe complex phenomena such as the inheritance of quantitative traits, the effect of selection, the consequence of mating behaviour and others. However, within such models polygenes are usually dealt with as a whole, whereas the actual genes remain in what has been defined as a 'statistical fog' [2]. Within this framework, the problem of understanding the molecular nature of quantitative trait variation would have remained unsolved.

The first steps toward resolving this puzzle were based on studies carried out during the first half of the 20th century; these showed that genes with a major effect on quantitative traits do exist and can be experimentally mapped on chromosomes by evaluating the correlation between the quantitative trait value and the allelic states at linked genetic markers [3,4]. This led to the definition of a quantitative trait locus (QTL) as a genetic locus where functionally different alleles segregate and cause significant effects on a quantitative trait. The findings of QTL

studies completed to date indicate an L-shaped distribution of QTL effects (i.e. most QTLs have a small effect and only a few show a strong effect) [5], thus enabling the identification of QTLs with a major effect on the phenotype.

Currently, QTL mapping is a standard procedure in quantitative genetics [6]. QTL mapping usually begins with the collection of genotypic (based on molecular markers) and phenotypic data from a segregating population, followed by statistical analysis to reveal all possible marker loci where allelic variation correlates with the phenotype. Because this procedure only allows for an approximate mapping of the QTL, it is usually referred to as primary (or coarse) QTL mapping. Procedures and strategies for primary QTL mapping are well established [7] and will not be considered here.

More recent technical progress in the area of molecular biology and genomics have made the cloning of QTLs [i.e. the identification of the DNA sequences (coding or non-coding) responsible for QTLs] possible. Here, we present a critical appraisal of the results obtained in this field in plants and discuss the perspectives, with emphasis on several major limitations and promising novel approaches. A literature survey shows that although ~150 research papers reporting original QTL data are published yearly (average of 2000–2004, considering *Arabidopsis*, soybean, rice, sorghum, maize, barley and wheat), only a handful of studies have reported the cloning of QTLs (Table 1). Besides plants, QTL cloning is rapidly advancing in humans and livestock [8,9] as well as in model species such as yeast, *Drosophila*, mouse and rat [10–12].

Positional cloning of QTLs

Among the handful of QTLs isolated in plants to date, the majority have been cloned via positional cloning. Positional cloning requires several steps (Figure 1) to enable us to assign a QTL to the shortest possible genetic interval (QTL fine genetic mapping) and to identify the corresponding interval on the DNA sequence (QTL physical mapping) where candidate genes are selected for evaluation. The increase in mapping resolution required by QTL positional cloning is substantial because after primary mapping a QTL is positioned within a chromosome interval of ~10–30 cM, which usually includes several hundred genes (Box 1). Eventually, independent proof is required to validate the role of the identified allelic polymorphism on the observed phenotypic effect.

Corresponding author: Tuberosa, R. (roberto.tuberosa@unibo.it).

Available online 10 May 2005

Table 1. Summary of the main characteristics of the QTLs cloned in plants

Species	Trait	QTL	Gene	Function	Molecular identification	Candidate gene ^a	R ² (%) ^b	Plants (no.) ^c	Resolution (kb) ^d	ORF (no.) ^e	Identification of QTN	Functional proof	Refs
Arabidopsis	Flowering time	ED1	CRY2	Crypt.	Pos. cloning	Yes (L)	28–56	1822	46	15	Amino acid substitution	Transformation	[71]
		FLW	FLM	TF	Pos. cloning	Yes (E)	27	NA	138	38	38	Deletion of whole gene	Transformation
	Gluc. structure	GS-elong	MAM	MAM synthase	Pos. cloning	Yes (E)	NA	4600	NA	NA	Nucleotide and gene indels	No	[82]
Maize	Root morphology	BRX	BRX	TF	Pos. cloning	No	80	860	45	10	Premature stop codon	Transformation	[83]
		Tb1	Tb1	TF	Transp. tagging	Yes (E)	17–31	NA	NA	NA	No, possibly regulatory	Complementation	[33,42]
Rice	Heading time	Hd1	Se1	TF	Pos. cloning	Yes (L)	67	1505	12	2	No	Transformation	[84]
	Heading time	Hd3a	Hd3a	Unknown	Pos. cloning	Yes (L)	NA	2207	20	4	No	Transformation	[23]
	Heading time	Hd6	αCK2	Protein kinase	Pos. cloning	No	NA	2807	26	1	Premature stop codon	Transformation	[30]
	Heading time	Ehd1	Ehd1	B-type response regulator	Pos. cloning	No	NA	> 2500	16	3	Amino acid substitution	Transformation	[85]
Tomato	Fruit sugar content	Brix9-2-5	Lin5	Invertase	Pos. cloning	Yes (L)	NA	7000	0.5	1	Amino acid substitution	Complementation	[86,87]
		Ovate	Ovate	Unknown	Pos. cloning	No	48–67	3000	55	8	Premature stop codon	Transformation	[88]
	Fruit weight	fw2.2	ORFX	Unknown	Pos. cloning	No	30	3472	92	4	Unknown regulatory variant	Transformation	[58,89]

Abbreviation: crypt., cryptochrome; gluc. structure, glucosinolate structure; MAM synthase, methylthioalkylmalate synthase; NA, not applicable or not available; ORF, open reading frame; pos. cloning, positional cloning; QTL, quantitative trait locus; QTN, quantitative trait nucleotide; TF, transcription factor; transp. tagging, transposon tagging.

^aEvidence for candidate gene: (E) indicates early evidence, after primary QTL analysis; (L) indicates late evidence, after physical mapping and/or sequencing.

^bProportion of phenotypic variance explained by the QTL in the primary cross.

^cDimension of the population used for fine mapping.

^dDNA physical interval completely linked with the QTL.

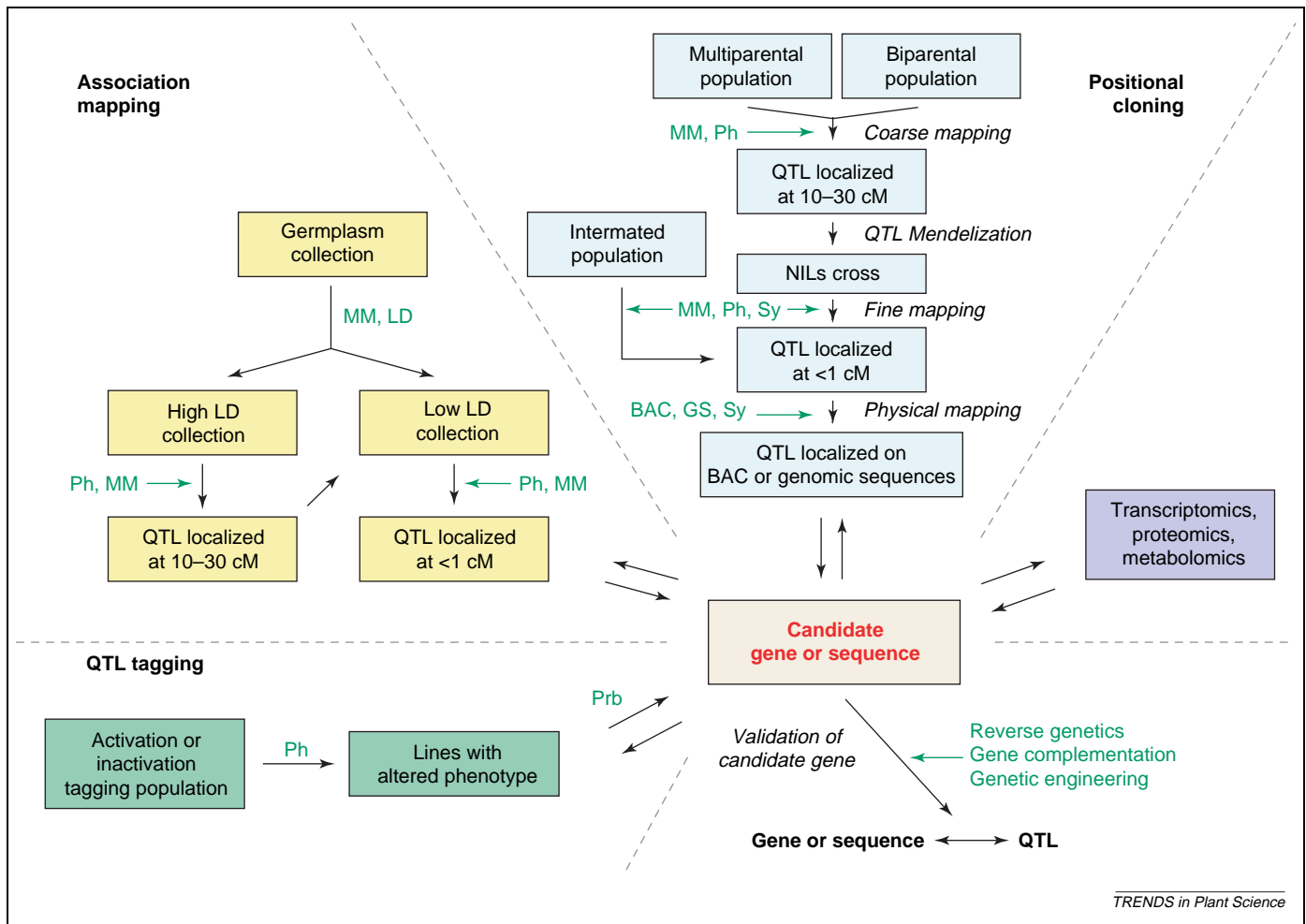
^eNumber of ORFs completely linked with the QTL.

Types of genetic material suitable for QTL fine mapping

A widely adopted strategy to estimate the position and effect of a coarsely mapped QTL more accurately is to create a new experimental population by crossing nearly isogenic lines (NILs) that differ only in the allelic constitution at the short chromosome segment (usually varying from ~10 to 30 cM in length) harbouring the QTL (QTL-NILs). In such a population, because of the absence of other segregating QTLs, the target QTL becomes the major genetic source of variation, and the phenotypic means of the QTL genotypic classes (+/+, -/- and, when present, -/+) can be statistically differentiated and genotypes recognized accordingly. Appropriate replication and/or progeny testing are generally implemented based upon the heritability of the trait considered. Under such conditions, the QTL is considered Mendelized [13], and cM distances between a QTL and the nearby molecular markers can be estimated more precisely.

Beginning with the same population in which primary mapping was carried out, QTL-NILs can be produced by (i) marker-assisted backcross introgression (i.e. substitution) of one QTL allele into one or both parental genetic backgrounds or (ii) iteratively identifying and selfing individuals that are heterozygous at the QTL region. QTL-NILs can also be efficiently identified within introgression libraries (ILs) (i.e. collections of lines where each line is isogenic to a background parental line, with the exception of a single short chromosome segment introgressed from a donor) [14]. Within an IL, the donor genome should be completely represented among the different IL lines, therefore, potentially a QTL-NIL exists for any segregating QTL. Remarkably, the same tomato IL [15] provided the source of QTL-NILs used for cloning three tomato QTLs. NILs suitable for positional cloning can also be produced using the advanced backcross QTL analysis (ABQA) method, which combines backcrossing chromosome segments from a wild accession within an elite line coupled with some level of phenotypic selection against extreme phenotypes [16]. Crucial aspects to be considered are the time and effort required for developing collections of IL or ABQA lines, as well as the limited genetic variability as a result of using only two parental lines: no matter how carefully the parental lines are chosen, the collections will only segregate for a fraction of the many more QTLs segregating for the same trait in other populations. This important limitation can be partially overcome through the use of multiparental intercrossed populations [17,18]. These populations are generated by crossing a carefully chosen set of parental lines capturing much of the genetic variation of the species, followed by several cycles of intermating. This approach should increase the efficiency of QTL mapping both in terms of detection (segregation is expected at many loci) and genetic resolution (many rounds of meiosis). A substantial increase in genetic resolution can also be obtained by intercrossing standard biparental populations [19].

During the fine-mapping step, the resolution of the target QTL in two or more linked loci can bring positional cloning projects to an end when the proportion of phenotypic variability explained by each QTL is too small to be revealed with a realistically manageable number of



TRENDS in Plant Science

Figure 1. Flow-chart depicting the molecular dissection of quantitative traits by positional cloning, association mapping and QTL-tagging. Boxes indicate starting materials or major milestones. Major experimental processes are indicated in black italic font. For each experimental process, relevant genetic, molecular or analytical tools are indicated in green font. Abbreviations: BAC, library of bacterial artificial chromosomes; GS, genomic sequence; LD, data on linkage disequilibrium; MM, molecular markers; NILs, nearly isogenic lines; Ph, phenotyping; Prb, probing with tagging agent, for example, transposon; Str, data on population structure; Sy, synteny. Reverse genetics includes transposon and T-tagging, activation tagging, TILLING and RNAi.

replications. QTL clusters have been observed in plants [20–22]; however, cloning was accomplished when one of the linked QTLs retained most of the effect [23,24].

Molecular genotyping

The recruitment of polymorphic markers required for fine mapping a QTL is fairly simple for *Arabidopsis* and rice because the whole genome has been sequenced, and also for species such as maize or tomato for which genomic sequencing is under way or information is available in terms of, for example, ESTs and BAC ends (See Glossary [25,26]). However, in species for which detailed sequence

information is not available or cannot be deduced from syntenic relatives, many molecular markers (e.g. amplified fragment length polymorphisms) need to be screened in genotypes contrasted at the target region (e.g. pair of QTL-NILs). Synteny with *Arabidopsis*, *Medicago* or *Lotus*, and rice should assist in identifying additional markers for cloning projects within *Brassicaceae*, legumes and cereals, respectively [27]. With regards to genotyping techniques, microarray-based platforms appear to be particularly promising for high-throughput identification of polymorphisms [single nucleotide polymorphisms (SNPs) and indels] at thousands of loci [28,29].

Box 1. How many genes underlie a primary QTL?

It is impossible to identify the gene subtending a specific QTL after its mapping by primary analysis because of the poor resolution of the analysis itself. Primary analysis usually maps a QTL within a chromosome region (known as QTL supporting interval) of ~10–30 cM. Using current estimates of the total gene number and the genetic and physical length of *Arabidopsis* and maize genomes (available at <http://www.arabidopsis.org/> and <http://www.maizegdb.org/>), it can be calculated that a 10-cM chromosome interval on average corresponds to ~2.1 Mb and 440 genes in *Arabidopsis* or to ~12.4 Mb and 310 genes in maize.

Physical mapping and the identification of candidate sequences

When the genetic resolution approaches the cM level, the markers closest to the QTL are used for anchoring the genetic map to the physical map (i.e. the genomic sequence or a BAC contig covering the QTL region). Early transfer of the information to the physical map enables the efficient generation of new single-copy markers useful for refining the genetic mapping and for searching candidate genes. Even if only a BAC contig is available, sequenced BAC

Glossary

BAC contig: a contiguous set of overlapping BAC clones.

BAC ends: portions of a BAC genomic insert that are near the cloning site of the vector. BAC ends can be easily sequenced and used to anchor the BAC clone to other sequences or to a genetic map.

BAC (bacterial artificial chromosome): a plasmid vector capable of 100–150 kb inserts.

Candidate gene: a gene that based on its map position and/or its nucleotide sequence could be responsible for a given phenotype.

cM (centiMorgan): indicates the distance of two loci on a chromosome based on the observed or estimated frequency of crossovers over the total number of scored meioses. For low values, one cM corresponds to 1% of gametes carrying a crossover chromosome.

EST (expressed sequenced tag): short (usually <500 bp) sequence, usually obtained by raw sequencing of a random cDNA. Collections of ESTs provide a quick way to represent a substantial portion of the gene complement of a genome.

Haplotype: a combination of alleles at different loci on the same chromosome segment.

Kb/cM rate: the average chromosome distance (in base pairs of DNA) per unit of genetic distance (cM). It can refer to a general mean of a species (computed as genome dimension in kb/total genetic map) or to a specific chromosome region.

LD (linkage disequilibrium): the level of non-random assortment of alleles at different loci.

Molecular marker: a locus whose genotype can be inferred by a molecular assay based on more or less direct analysis of the DNA sequence.

Phenotypic variance: a statistical index of the variability of the population. It includes two major components: variability produced by genetic segregation and interaction of QTLs and variability produced by random environmental effects.

QTN (quantitative trait nucleotide): the DNA sequence polymorphism responsible for the QTL effect.

RNAi (RNA interference): a biological process (and a reverse genetics technique) causing post-transcriptional inhibition of the expression of a target gene by the action of small (21–26 nucleotides) RNA molecules.

Segregating population: a population of individuals differing in allelic constitution at one or more genetic loci.

SNP (single nucleotide polymorphism): a point mutation that can be targeted by molecular techniques and can thus be exploited as a molecular marker.

Somaclonal variation: the *de novo* variation observed in plants regenerated from tissue culture. It can be caused by transposon activity, late replication of heterochromatin, gene amplification, mitotic crossing-over and/or changes in methylation that occur during the *in vitro* phase.

Syteny: conservation of genome organization (including the linear order of genes on chromosomes, which is called collinearity) among phylogenetically related species.

TILLING (targeting induced local lesions in genomes): a reverse genetics technique that enables individuals carrying point mutations at a target gene of known sequence to be identified within a chemically mutagenized population.

ends can often be transformed in genetic markers, and low-pass shot-gun sequencing can provide a glimpse of local gene content. In this phase, bioinformatics provides an important contribution in terms of gene prediction and annotation, and exploitation of syntenic relationships.

Among the studies considered here, only two managed to reduce the number of genes co-segregating with the target QTL (Table 1) to one. In one case, *Hd6*, the QTN (i.e. the quantitative trait nucleotide polymorphism responsible for the QTL effect) was a nucleotide substitution in one of the two alleles that caused a premature stop codon [30]. When the physical region co-segregating with the QTL includes more than one gene (can be up to 38; [31]), candidates can either be identified via function prediction and selected for further testing (see *Cry2*, *FLM*, *Hd1* and *Hd3a* in Table 1) or they might not be clearly evident. When multiple coding sequences with no obvious candidate gene are identified, two possible options are to increase the mapping resolution or to test each open reading frame (ORF) functionally.

Validation of a candidate gene or sequence

The functional testing of a candidate gene(s) can be performed by overexpressing or down-regulating the target gene through genetic engineering or RNAi [32], or by genetic complementation of a known mutant [33]. If available within the species under investigation, reverse genetics tools such as T-DNA or transposon-tagged populations [34] and/or TILLING [35] can also be exploited. Compared with transposon tagging, TILLING and RNAi are appealing alternatives for their almost universal applicability and for providing subtle changes of gene functionality comparable to those observed naturally. Gene replacement, still in its infancy but already reported in rice [36], is the ultimate tool for validating candidate genes.

The validation of QTNs in non-coding regions is one of the current major challenges. Regulatory regions close to (e.g. promoters) or far (e.g. enhancers or silencers) from the regulated gene are likely to host sequence polymorphisms causing variation in quantitative phenotypes. It can also be predicted that QTLs will be found at microRNA loci and at regions controlling chromatin methylation and/or organization (e.g. folding). Furthermore, transposon insertions have already been shown to be responsible for changes in gene expression [37]. However, for most of these genomic features a structural characterization is still lacking, thus hindering the recognition of their role in the control of quantitative traits. Regulatory elements at several tens of kb from coding sequences have already been shown to act as QTLs [38–40], even though the causal QTNs have not yet been identified.

After considering all the above-mentioned aspects, it is clear that positional cloning of QTLs in plants remains a demanding and daunting undertaking. Furthermore, positional cloning has been limited exclusively to major QTLs because all the cloned QTLs showed an R^2 value higher than 15% in the primary genetic analysis (Table 1). R^2 values can be grossly under- or overestimated [41] because of statistical artefacts and because epistasis can modify the genetic effect of the target QTL when the genetic background changes [42], for instance during QTL-NIL production. Therefore, an independent evaluation of the QTL effect (e.g. by developing and phenotypically testing QTL-NILs [43]) is recommended before embarking on QTL positional cloning.

Cloning QTLs by association mapping

As an alternative to positional cloning, QTLs can be molecularly resolved through association mapping [44] (i.e. by identifying, within a set of genotypes such as germplasm accessions and cultivated varieties, a statistical association between allelic variants at marker or candidate loci and the mean of the analysed trait). The analysis evaluates the trait mean change caused by the substitution of one allele with another. For QTL cloning in plants, the interest lies in (i) the possibility of finding chromosome regions important for controlling quantitative traits without the costly and time-consuming production of large experimental populations [45], (ii) the potentially high genetic resolution provided by the many meiotic events that occurred during past generations, and (iii) the

possibility of surveying many functionally diverse alleles per locus.

A major factor to be considered in association mapping is the level of linkage disequilibrium (LD) among the tested accessions. In plants, the extensive LD analyses conducted in *Arabidopsis* and maize [46] have indicated that whereas LD persists over hundreds of kb in *Arabidopsis*, in maize, LD decays after a few kb, although it can extend significantly farther in collections of elite germplasm [46]. With high LD values, marker-trait association can theoretically be revealed with a manageable number of molecular markers. In this case, the expected mapping resolution will only be sufficient for the discovery and coarse mapping of the QTL. However, when testing germplasm panels with low LD, the diagnostic power of a single marker will only extend a short way and thus a prohibitively high number of markers would be required for a whole genome scan. In this situation, association mapping can still be used to fine map the QTL at the gene level after the QTL is positioned using standard mapping procedures. Based on this, it is conceivable that different sets of genotypes, characterized by high or low LD, can be assembled and used for QTL discovery or candidate gene validation, respectively, as has been suggested for human genetics [47]. Population structure (i.e. the possible presence of hidden subgroups, because of e.g. relatedness and selection, with an unequal distribution of alleles) might influence the efficacy of this approach by causing spurious trait-marker associations [48].

EcoTILLING [49] is a powerful approach for identifying different haplotypes (combinations of allelic variants) at target loci and making them available for association mapping, enabling the identification of virtually all SNPs and small insertion or deletions within a ~1-kb window in a set of genotypes at a fraction of the sequencing cost. However, the necessity of also screening regulatory regions that are often distant from the effector genes indicates that selecting the candidate sequences to be tested for association mapping is not a trivial task if the genomic scan aims to be comprehensive. Examples of identifying associations between haplotype variation at a candidate gene and a quantitative trait have been reported in *Arabidopsis* [50], *Brassica* [51,52], potato [53] and maize [54–56].

The identification of a statistically significant association between haplotype variation at a candidate gene or sequence and a quantitative phenotype should be followed by validation experiments similar to those used in the positional cloning approach previously described.

Functional genomics and QTL cloning

The use of functional genomics is contributing to many aspects of QTL analysis and cloning. Transcriptional profiling between contrasting QTL genotypes can quickly provide a list of genes differentially expressed; subsequently, those genes functionally related to the target trait and mapping at the QTL region can be selected as candidates [57]. Unfortunately, the number of QTLs cloned to date in plants is too small to test the validity of this approach. Indeed, when the QTL has been shown to involve a difference in gene expression level between

alleles, the difference was either too low (approximately twofold [33]) or showed too strong a spatial and/or temporal pattern [58] to enable them to be identified using standard microarray-based transcriptome analysis. Other profiling platforms, such as SAGE (serial analysis of gene expression [59]) and MPSS (massively parallel signature sequencing [60]) are better suited to detecting subtle differences in gene expression. Transcript profiling can reach the sub-tissue level of resolution if carried out in combination with laser-capture microscopy [61].

The expression profiling of a mapping population at the mRNA or protein level enables us to treat the level of expression of a single gene as a quantitative trait and to dissect its genetic control by QTL analysis [5,62–64]. The loci controlling the level of gene expression have variously been named transcript quantity loci (TQLs), expression QTLs (eQTLs) or protein quantity loci (PQLs) [64,65]. Correspondence between eQTLs and/or PQLs and candidate genes with QTLs for morpho-physiological traits has already been observed in small- or medium-scale experiments [66–68]. Microarray-based studies have mapped eQTLs at the same location of the gene whose expression was measured, thus indicating a role for *cis*-regulatory allelic variation, and also at distant chromosome positions [65,69]. The same studies highlighted the presence of eQTL ‘hot spots’ (i.e. chromosome regions apparently responsible for controlling the simultaneous expression of many genes).

Adding a new framework: mutations at QTLs and QTL tagging

Donald Robertson [70] suggested that qualitative mutant alleles and wild-type alleles at loci affecting quantitative traits are the extremes of a possible range of effects, with QTLs resulting from the segregation of naturally available alleles with milder effects. Robertson’s hypothesis was confirmed in those cases where a mutant was available for the gene subtending the target QTL [42,71]. Along this line, it was recently argued that mutagenesis could be more efficient for dissecting the genetic basis of quantitative traits than is QTL analysis, which only provides ‘accidents of history’ allelic variants [72]. A direct method for identifying such genes would be to use a tagging (insertional) approach. Such a framework would require the phenotypic screening of an insertional mutagenized population for the target quantitative trait to identify those lines with a phenotypic mean value outside a predicted range because of environmental effects [70,73]. The complete screening experiment would involve a manageable number of plants (e.g. up to 20 000–30 000) if multiple insertion systems are used and several quantitative traits are concurrently evaluated [73]. The gene functionally modified or inactivated by the insertional agent could be rescued using standard molecular procedures. Following a similar approach, QTL tagging has already been successfully accomplished in *Magnaporthe* [74], the causal agent of rice blast, and in *Drosophila* [75]. In plants, QTL tagging could be carried out using several different approaches, based on T-DNA as well as DNA-transposons and retrotransposons. However, systems relying on callus cultures (e.g. activation of rice

TOS-17 retrotransposon [76]) should be considered with caution because of the occurrence of somaclonal variation [77], which can potentially alter any quantitative trait and therefore hinder the identification of the tagged QTL. Instead, other interesting resources are the *Ac-Ds*-based insertional populations developed in rice [78]; following the introduction of heterologous transposons, the majority of mutational events are created by new transposition activity. In maize, a *Mu*-based insertional population has been developed in a non-segregating genetic background [79]; within such a population, most of the quantitative variability can be attributed to the segregation of the tagged QTLs.

Increasingly important role of candidate genes

Classically, a link between a gene and a quantitative trait can be hypothesized based on linkage information (all genes co-segregating with a QTL are positional candidate genes) or commonality between the quantitative trait physiology and the biochemical function of the gene (functional candidate gene) [80], or both. For instance, completion of genome sequences and improved bioinformatics should facilitate *in silico* cross-matching of candidate sequences with QTLs in programmes of positional cloning or association mapping. The creation of more powerful bioinformatic tools for gene annotation should facilitate the choice of functional candidates among and outside the positional candidate genes [81]. In addition, a better understanding of the mechanisms behind the regulation of gene expression should extend the concept of candidate gene to include *cis*-acting regulatory sequences. Therefore, in the future, it is conceivable that QTL cloning will increasingly rely on candidate gene information and that this will be made possible by exploiting the available reverse genetics tools (Figure 1).

Conclusions

The QTLs cloned to date in plants are likely to represent a biased sample of those governing the variability of target traits; only major QTLs, mostly identified in wide crosses (e.g. *indica* × *japonica* rice subspecies, teosinte × cultivated maize and wild × cultivated tomato), have been successfully targeted. However, the constant improvement of the molecular platforms, new types of genetic materials, progress in bioinformatics and the increasing availability of tools and platforms for functionally testing candidate genes should facilitate QTL cloning and offer the opportunity of targeting QTLs other than those with a major effect. Quantitative approaches to dissect mutants genetically will also have to be extended if more subtle regulators of complex phenotypes are to be uncovered.

From a breeding standpoint, we need to continue investigating natural variation to unlock the allelic richness present in germplasm collections, which should enable us to use marker-assisted selection and genetic engineering more effectively to introduce valuable alleles in crops. At the same time, QTL cloning should improve our understanding on how nature shaped genetic variability during adaptive evolution.

References

- Mather, K. (1941) Variation and selection of polygenic characters. *J. Genet.* 41, 159–193
- Mauricio, R. (2001) Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology. *Nat. Rev. Genet.* 2, 370–381
- Sax, K. (1923) The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics* 8, 552–560
- Thoday, J.M. (1961) Location of polygenes. *Nature* 191, 368–370
- Brem, R.B. and Kruglyak, L. (2005) The landscape of genetic complexity across 5,700 gene expression traits in yeast. *Proc. Natl. Acad. Sci. U. S. A.* 102, 1572–1577
- Lynch, M. and Walsh, B., eds (1998) *Genetics and Analysis of Quantitative Traits*, Sinauer Associates
- Hackett, C.A. (2002) Statistical methods for QTL mapping in cereals. *Plant Mol. Biol.* 48, 585–599
- Korstanje, R. and Paigen, B. (2002) From QTL to gene: the harvest begins. *Nat. Genet.* 31, 235–236
- Andersson, L. and Georges, M. (2004) Domestic-animal genomics: deciphering the genetics of complex traits. *Nat. Rev. Genet.* 5, 202–212
- Glazier, A.M. *et al.* (2002) Finding genes that underlie complex traits. *Science* 298, 2345–2349
- Mackay, T.F.C. (2004) The genetic architecture of quantitative traits: lessons from *Drosophila*. *Curr. Opin. Genet. Dev.* 14, 253–257
- Steinmetz, L.M. *et al.* (2002) Dissecting the architecture of a quantitative trait locus in yeast. *Nature* 416, 326–330
- Alonso-Blanco, C. and Koornneef, M. (2000) Naturally occurring variation in *Arabidopsis*: an underexploited resource for plant genetics. *Trends Plant Sci.* 5, 22–29
- Zamir, D. (2001) Improving plant breeding with exotic genetic libraries. *Nat. Rev. Genet.* 2, 983–989
- Eshed, Y. and Zamir, D. (1994) A genomic library of *Lycopersicon pennellii* in *L. esculentum*: a tool for fine-mapping of genes. *Euphytica* 79, 175–179
- Tanksley, S.D. and Nelson, J.C. (1996) Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor. Appl. Genet.* 92, 191–203
- The Complex Trait Consortium. (2004) The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nat. Genet.* 36, 1133–1137
- Mott, R. *et al.* (2000) A method for fine mapping quantitative trait loci in outbred animal stocks. *Proc. Natl. Acad. Sci. U. S. A.* 97, 12649–12654
- Lee, M. *et al.* (2002) Expanding the genetic map of maize with the intermated B73 × Mo17 (IBM) population. *Plant Mol. Biol.* 48, 453–461
- Chen, K.Y. and Tanksley, S.D. (2004) High-resolution mapping and functional analysis of *se2.1*: a major stigma exertion quantitative trait locus associated with the evolution from allogamy to autogamy in the genus *Lycopersicon*. *Genetics* 168, 1563–1573
- Tuberosa, R. *et al.* (2002) Mapping QTLs regulating morphophysiological traits and yield: case studies, shortcomings and perspectives in drought-stressed maize. *Ann. Bot. (Lond.)* 89, 941–963
- Khavkin, E. and Coe, E. (1997) Mapped genomic locations for developmental functions and QTLs reflect concerted groups in maize (*Zea mays* L.). *Theor. Appl. Genet.* 95, 343–352
- Kojima, S. *et al.* (2002) *Hd3a*, a rice ortholog of the *Arabidopsis FT* gene, promotes transition to flowering downstream of *Hd1* under short-day conditions. *Plant Cell Physiol.* 43, 1096–1105
- Fridman, E. *et al.* (2002) Two tightly linked QTLs modify tomato sugar content via different physiological pathways. *Mol. Genet. Genomics* 266, 821–826
- Martienssen, R.A. *et al.* (2004) Sequencing the maize genome. *Curr. Opin. Plant Biol.* 7, 102–107
- Van der Hoeven, R. *et al.* (2002) Deductions about the number, organization, and evolution of genes in the tomato genome based on analysis of a large expressed sequence tag collection and selective genomic sequencing. *Plant Cell* 14, 1441–1456
- Schmidt, R. (2000) Synteny: recent advances and future prospects. *Curr. Opin. Plant Biol.* 3, 97–102
- Borevitz, J.O. *et al.* (2003) Large-scale identification of single-feature polymorphisms in complex genomes. *Genome Res.* 13, 513–523

- 29 Hardenbol, P. *et al.* (2003) Multiplexed genotyping with sequence-tagged molecular inversion probes. *Nat. Biotechnol.* 21, 673–678
- 30 Takahashi, Y. *et al.* (2001) *Hd6*, a rice quantitative trait locus involved in photoperiod sensitivity, encodes the α -subunit of protein kinase CK2. *Proc. Natl. Acad. Sci. U. S. A.* 98, 7922–7927
- 31 Werner, J.D. *et al.* (2005) Quantitative trait locus mapping and DNA array hybridization identify an *FLM* deletion as a cause for natural flowering-time variation. *Proc. Natl. Acad. Sci. U. S. A.* 102, 2460–2465
- 32 Waterhouse, P.M. and Helliwell, C.A. (2003) Exploring plant genomes by RNA-induced gene silencing. *Nat. Rev. Genet.* 4, 29–38
- 33 Doebley, J. *et al.* (1997) The evolution of apical dominance in maize. *Nature* 386, 485–488
- 34 Maes, T. *et al.* (1999) Plant tagnology. *Trends Plant Sci.* 4, 90–96
- 35 McCallum, C.M. *et al.* (2000) Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. *Plant Physiol.* 123, 439–442
- 36 Iida, S. and Terada, R. (2004) A tale of two integrations, transgene and T-DNA: gene targeting by homologous recombination in rice. *Curr. Opin. Biotechnol.* 15, 132–138
- 37 Kashkush, K. *et al.* (2003) Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat. Genet.* 33, 102–106
- 38 Clark, R.M. *et al.* (2004) Pattern of diversity in the genomic region near the maize domestication gene *tb1*. *Proc. Natl. Acad. Sci. U. S. A.* 101, 700–707
- 39 Salvi, S. *et al.* (2002) Towards positional cloning of *Vgt1*, a QTL controlling the transition from the vegetative to the reproductive phase in maize. *Plant Mol. Biol.* 48, 601–613
- 40 Salvi, S. *et al.* (2003) Progress in the positional cloning of *Vgt1*, a QTL controlling flowering time in maize. In *Proceedings of 57th Corn and Sorghum and 32nd Soybean Seed Research Conference*, American Seed Trait Association, Chicago, IL, USA
- 41 Beavis, W.D. (1994) The power and deceit of QTL experiments: lessons from comparative QTL studies. In *Proc. 49th Ann. Corn and Sorghum Res. Conf. American Seed Trait Association*, pp. 250–266, American Seed Trait Association, Washington, D.C., USA
- 42 Doebley, J. *et al.* (1995) *Teosinte branched1* and the origin of maize: evidence for epistasis and the evolution of dominance. *Genetics* 141, 333–346
- 43 Landi, P. *et al.* (2005) Validation and characterization of a major QTL affecting leaf ABA concentration in maize. *Mol. Breed.* 15, 291–303
- 44 Cardon, L.R. and Bell, J.I. (2001) Association study designs for complex diseases. *Nat. Rev. Genet.* 2, 91–99
- 45 Morgante, M. and Salamini, F. (2003) From plant genomics to breeding practice. *Curr. Opin. Biotechnol.* 14, 214–219
- 46 Flint-Garcia, S.A. *et al.* (2003) Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* 54, 357–374
- 47 Reich, D.E. *et al.* (2001) Linkage disequilibrium in the human genome. *Nature* 411, 199–204
- 48 Pritchard, J.K. *et al.* (2000) Association mapping in structured populations. *Am. J. Hum. Genet.* 67, 170–181
- 49 Comai, L. *et al.* (2004) Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. *Plant J.* 37, 778–786
- 50 Olsen, K.M. *et al.* (2004) Linkage disequilibrium mapping of *Arabidopsis* *CRY2* flowering time alleles. *Genetics* 167, 1361–1369
- 51 Osterberg, M.K. *et al.* (2002) Naturally occurring indel variation in the *Brassica nigra* *COL1* gene is associated with variation in flowering time. *Genetics* 161, 299–306
- 52 Gupta, V. *et al.* (2004) Molecular tagging of erucic acid trait in oilseed mustard (*Brassica juncea*) by QTL mapping and single nucleotide polymorphisms in *FAE1* gene. *Theor. Appl. Genet.* 108, 743–749
- 53 Simko, I. *et al.* (2004) Linkage disequilibrium mapping of a *Verticillium dahliae* resistance quantitative trait locus in tetraploid potato (*Solanum tuberosum*) through a candidate gene approach. *Theor. Appl. Genet.* 108, 217–224
- 54 Thornsberry, J.M. *et al.* (2001) *Dwarf8* polymorphisms associate with variation in flowering. *Nat. Genet.* 28, 286–289
- 55 Guillet-Claude, C. *et al.* (2004) Nucleotide diversity of the *ZmPox3* maize peroxidase gene: relationships between a MITE insertion in exon 2 and variation in forage maize digestibility. *BMC Genet.* 5, 19
- 56 Wilson, L.M. *et al.* (2004) Dissection of maize kernel composition and starch production by candidate gene association. *Plant Cell* 16, 2719–2733
- 57 Wayne, M.L. and McIntyre, L.M. (2002) Combining mapping and arraying: an approach to candidate gene identification. *Proc. Natl. Acad. Sci. U. S. A.* 99, 14903–14906
- 58 Cong, B. *et al.* (2002) Natural alleles at a tomato fruit size quantitative trait locus differ by heterochronic regulatory mutations. *Proc. Natl. Acad. Sci. U. S. A.* 99, 13606–13611
- 59 Gowda, M. *et al.* (2004) Robust-LongSAGE (RL-SAGE): a substantially improved LongSAGE method for gene discovery and transcriptome analysis. *Plant Physiol.* 134, 890–897
- 60 Brenner, S. *et al.* (2000) Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat. Biotechnol.* 18, 630–634
- 61 Schnable, P.S. *et al.* (2004) Global expression profiling applied to plant development. *Curr. Opin. Plant Biol.* 7, 50–56
- 62 Jansen, R.C. and Nap, J.P. (2001) Genetical genomics: the added value from segregation. *Trends Genet.* 17, 388–391
- 63 Tuberosa, R. *et al.* (2003) Searching for quantitative trait loci controlling root traits in maize: a critical appraisal. *Plant Soil* 255, 35–54
- 64 Damerval, C. *et al.* (1994) Quantitative trait loci underlying gene product variation: a novel perspective for analyzing regulation of genome expression. *Genetics* 137, 289–301
- 65 Schadt, E. *et al.* (2003) Genetics of gene expression surveyed in maize, mouse and man. *Nature* 422, 297–302
- 66 de Vienne, D. *et al.* (1999) Genetics of proteome variation for QTL characterization: application to drought stress responses in maize. *J. Exp. Bot.* 50, 303–309
- 67 Francia, E. *et al.* (2003) Two loci on chromosome 5H determine low-temperature tolerance in a 'Nure' (winter) × 'Tremois' (spring) barley map. *Theor. Appl. Genet.* 108, 670–680
- 68 Guillaumie, S. *et al.* (2004) Colocation between a gene encoding the bZip factor *SPA* and an eQTL for a high-molecular-weight glutenin subunit in wheat (*Triticum aestivum*). *Genome* 47, 705–713
- 69 Brem, R.B. *et al.* (2002) Genetic dissection of transcriptional regulation in budding yeast. *Science* 296, 752–755
- 70 Robertson, D.S. (1985) A possible technique for isolating genic DNA for quantitative traits in plants. *J. Theor. Biol.* 117, 1–10
- 71 El-Din El-Assal, S. *et al.* (2001) A QTL for flowering time in *Arabidopsis* reveals a novel allele of *CRY2*. *Nat. Genet.* 29, 435–440
- 72 Nadeau, J.H. and Frankel, W.N. (2000) The roads from phenotypic variation to gene discovery: mutagenesis versus QTLs. *Nat. Genet.* 25, 381–384
- 73 Soller, M. and Beckmann, J.S. (1987) Cloning quantitative trait loci by insertional mutagenesis. *Theor. Appl. Genet.* 74, 369–378
- 74 Fujimoto, D. *et al.* (2002) Tagging quantitative loci controlling pathogenicity in *Magnaporthe grisea* by insertional mutagenesis. *Physiol. Mol. Plant Pathol.* 61, 77–88
- 75 Norga, K.K. *et al.* (2003) Quantitative analysis of bristle number in *Drosophila* mutants identifies genes involved in neural development. *Curr. Biol.* 13, 1388–1396
- 76 Hirochika, H. *et al.* (1996) Retrotransposons of rice involved in mutations induced by tissue culture. *Proc. Natl. Acad. Sci. U. S. A.* 93, 7783–7788
- 77 Kaeppler, S.M. and Phillips, R.L. (1993) Tissue culture-induced DNA methylation variation in maize. *Proc. Natl. Acad. Sci. U. S. A.* 90, 8773–8776
- 78 Jeon, J. and An, G. (2001) Gene tagging in rice: a high throughput system for functional genomics. *Plant Sci.* 161, 211–219
- 79 Settles, A.M. *et al.* (2004) Molecular analysis of high-copy insertion sites in maize. *Nucleic Acids Res.* 32, e54
- 80 Pflieger, S. *et al.* (2001) The candidate gene approach in plant genetics: a review. *Mol. Breed.* 7, 275–291
- 81 Harhay, G.P. and Keele, J.W. (2003) Positional candidate gene selection from livestock EST databases using Gene Ontology. *Bioinformatics* 19, 249–255
- 82 Kroymann, J. *et al.* (2003) Evolutionary dynamics of an *Arabidopsis* insect resistance quantitative trait locus. *Proc. Natl. Acad. Sci. U. S. A.* 100(Suppl. 2), 14587–14592
- 83 Mouchel, C.F. *et al.* (2004) Natural genetic variation in *Arabidopsis* identifies *BREVIS RADIX*, a novel regulator of cell proliferation and elongation in the root. *Genes Dev.* 18, 700–714
- 84 Yano, M. *et al.* (2000) *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. *Plant Cell* 12, 2473–2484

- 85 Doi, K. *et al.* (2004) *Ehd1*, a B-type response regulator in rice, confers short-day promotion of flowering and controls *FT*-like gene expression independently of *Hd1*. *Genes Dev.* 18, 926–936
- 86 Fridman, E. *et al.* (2000) A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. *Proc. Natl. Acad. Sci. U. S. A.* 97, 4718–4723
- 87 Fridman, E. *et al.* (2004) Zooming in on a quantitative trait for tomato yield using interspecific introgressions. *Science* 305, 1786–1789
- 88 Liu, J. *et al.* (2002) A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. *Proc. Natl. Acad. Sci. U. S. A.* 99, 13302–13306
- 89 Frary, A. *et al.* (2000) *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. *Science* 289, 85–88

Elsevier celebrates two anniversaries with gift to university libraries in the developing world

In 1580, the Elzevir family began their printing and bookselling business in the Netherlands, publishing works by scholars such as John Locke, Galileo Galilei and Hugo Grotius. On 4 March 1880, Jacobus George Robbers founded the modern Elsevier company intending, just like the original Elzevir family, to reproduce fine editions of literary classics for the edification of others who shared his passion, other 'Elzevirians'. Robbers co-opted the Elzevir family's old printer's mark, visually stamping the new Elsevier products with a classic old symbol of the symbiotic relationship between publisher and scholar. Elsevier has since become a leader in the dissemination of scientific, technical and medical (STM) information, building a reputation for excellence in publishing, new product innovation and commitment to its STM communities.

In celebration of the House of Elzevir's 425th anniversary and the 125th anniversary of the modern Elsevier company, Elsevier will donate books to 10 university libraries in the developing world. Entitled 'A Book in Your Name', each of the 6 700 Elsevier employees worldwide has been invited to select one of the chosen libraries to receive a book donated by Elsevier. The core gift collection contains the company's most important and widely used STM publications including *Gray's Anatomy*, *Dorland's Illustrated Medical Dictionary*, *Essential Medical Physiology*, *Cecil Essentials of Medicine*, *Mosby's Medical, Nursing and Allied Health Dictionary*, *The Vaccine Book*, *Fundamentals of Neuroscience*, and *Myles Textbook for Midwives*.

The 10 beneficiary libraries are located in Africa, South America and Asia. They include the Library of the Sciences of the University of Sierra Leone; the library of the Muhimbili University College of Health Sciences of the University of Dar es Salaam, Tanzania; the library of the College of Medicine of the University of Malawi; and the libraries of the University of Zambia, Université du Mali, Universidade Eduardo Mondlane, Mozambique; Makerere University, Uganda; Universidad San Francisco de Quito, Ecuador; Universidad Francisco Marroquin, Guatemala; and the National Centre for Scientific and Technological Information (NACESTI), Vietnam.

Through 'A Book in Your Name', the 10 libraries will receive approximately 700 books at a retail value of approximately 1 million US dollars.

For more information, visit www.elsevier.com