Full text provided by www.sciencedirect.com

SCIENCE DIRECT®

Genomics tools for QTL analysis and gene discovery

Justin O Borevitz[1] and Joanne Chory[1,2]

In recent years, several new genomics resources and tools have become available that will greatly assist quantitative trait locus (QTL) mapping and cloning of the corresponding genes. Genome sequences, tens of thousands of molecular markers, microarrays, and knock-out collections are being applied to QTL mapping, facilitating the use of natural accessions for gene discovery.

Addresses
[1]Plant Biology Laboratory, [2]Howard Hughes Medical Institute, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, California 92037, USA
[2]e-mail: chory@salk.edu

Current Opinion in Plant Biology 2004, **7**:132–136

This review comes from a themed issue on Genome studies and molecular genetics
Edited by Joseph R Ecker and Doug Cook

1369-5266/$ – see front matter
© 2004 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.pbi.2004.01.011

Abbreviations
EST expressed sequence tag
HIF heterogeneous inbred family
LD linkage disequilibrium
NIL near isogenic line
QTL quantitative trait locus/loci
SFP single feature polymorphism
SNP single-nucleotide polymorphism

Introduction

Quantitative trait locus (QTL) mapping has been in wide use for nearly two decades during which molecular markers have become available in conjunction with interval mapping methods [1]. The goal of QTL mapping is to determine the loci that are responsible for variation in complex, quantitative traits. In some situations, determination of the number, location and the interaction of these loci is the ultimate goal; often, however, the identification of the actual genes and their functions are of interest. For example, breeding studies attempt to identify the loci that improve crop yield or quality, and then to bring the favorable alleles together into elite lines. Understanding of the response of QTL in different environments or genetic backgrounds can lead to the development of improved crop varieties through traditional breeding. If the genes underlying the QTL are known (i.e. the QTL have been 'cloned'), then transgenic approaches can also be used to directly introduce beneficial alleles across wide species boundaries. In evolutionary studies, QTL define the genetic architecture of traits that are related to fitness and that differ between recently derived species [2]; however, knowledge of the actual genes allows for studies of molecular evolution. Studies of crop varieties and their wild progenitors have been effective in identifying large-effect QTL under artificial selection during crop domestication [3].

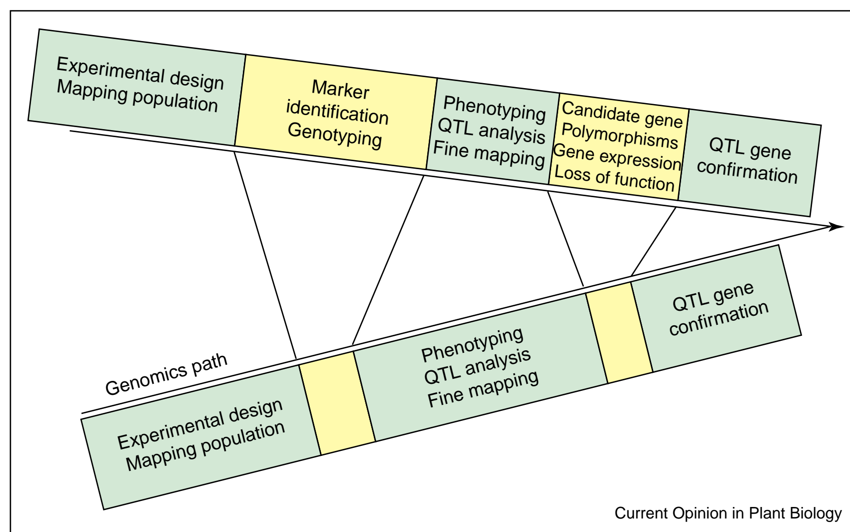
Until recently, QTL mapping was limited by the availability of molecular makers and the tediousness of their genotyping. Several high-throughput technologies that shift the burden of QTL mapping to gathering the phenotype information are now routine. In this review, we describe genomics tools and resources being used for QTL mapping (Figure 1) and cloning in *Arabidopsis thaliana*, including microarrays, which are used for both genotyping and gene expression analyses. Perhaps the most obvious genomic resource for QTL mapping is a complete genome sequence; but we also discuss what can be done with synteny and mapped expressed sequence tags (ESTs) when no complete sequence is available.

Molecular markers

Several studies have been published or are in progress that have, for all practical purposes, eliminated the need to identify new molecular markers in *Arabidopsis*. Cereon Genomics released 56 670 single-nucleotide polymorphism (SNP)/indel candidate polymorphisms from the 2–3X shotgun sequencing of the Landsberg *erecta Arabidopsi*s accession [4•]. Schmid *et al.* [5•] identified 8688 candidate SNP/indel polymorphisms from EST and sequence-tagged site (STS) reads of up to 12 accessions; in this case, the approximate allele frequencies are known for many polymorphisms. Magnus Nordborg plans to sequence up to 2000 fragments from a plate of 96 accessions for SNP discovery and linkage disequilibrium studies. At the time of writing, 15 388 polymorphisms are available from 824 fragments (M Norgborg, unpublished; http://walnut.usc.edu). The physical positions of 1267 traditional amplified length fragment polymorphisms (AFLPs) are also available [6].

Genotyping technologies, which allow these markers to be processed quickly, have also come on line rapidly. Several approaches require that fragments spanning an SNP be amplified, after which extension reactions interrogate the polymorphic base [7–9]. Several approaches differ mainly in the way that the alternative alleles are detected. Often, the SNP amplification reactions can be multiplexed; but usually, a maximum of 10 SNPs can be amplified per reaction. The individual marker price continues to decline as these methods become routine.

Figure 1



Genomics technologies have reduced several time-consuming steps (yellow boxes) in the identification of the genes that are responsible for quantitative traits. High-throughput genotyping techniques and large marker collections help at early stages, whereas gene expression and knock-out collections are important at late stages.

Array hybridization is another source of polymorphisms. Affymetrix-type high-density oligonucleotide arrays may contain millions of 25mer features. Each has the potential to identify a marker when the arrays are hybridized with labeled total genomic DNA [10••,11]. At present, data are available for more than 19 000 non-singleton single feature polymorphism (SFPs) from 14 accessions (http://naturalvariation.org/sfp). With SFPs, the actual base-pair change is not known; all that is known is that a particular 25mer is likely to have a change because of the differential hybridization of the genomic DNA from two parents. Thus, genomic DNA hybridized to expression arrays can be used as both a discovery and genotyping platform. This is an attractive method if many genotypes are required per sample. Spotted oligoarrays may be a less expensive alternative [12] but may suffer from lower reproducibility and batch-to-batch variation.

Analysis methods

Bulk segregant mapping is aptly suited to parallel genotyping methods. Array hybridization with DNA from pools of segregating lines, selected for alternative phenotypes, can quickly identify the location of large-effect mutations [10••]. Recently, we have extended this technique to quantitative traits by pooling lines that have extreme phenotypes (D Wolyn, JO Borevitz, J Chory, unpublished).

Linkage disequilibrium (LD) analysis in particular will benefit from high-resolution genotyping because several adjacent SNPs are needed to determine haplotypes. Hence, genomics has helped to realize LD as a tool for fine mapping of QTL [13]. Current studies aim to develop databases of high-resolution genotype information from a large collection of *Arabidopsis* accessions (http://walnut.usc.edu). This collection can be phenotyped for the trait of interest. With this information at hand, LD-mapping studies aim to associate quantitative phenotypes with haplotype information, a process known as '*in silico*' mapping [14]. The information provided by *in silico* maps, in conjunction with that from traditional QTL mapping studies, may provide a powerful way of quickly localizing QTL candidate genes. The extent of LD in *Arabidopsis* is estimated to be 50–250 kb [15], which limits the resolution for fine mapping but, conversely, makes it more likely that significant associations will be found in genome-wide scans. In maize, in which LD is on the scale of a few kilobases, associations can identify the underlying gene if properly controlled at other loci [16,17,18•,19].

Given plentiful markers and high-throughput genotyping technologies, QTL studies are limited by reliable phenotypic measures and multiple observations. Experimental design is therefore paramount. Every QTL experiment includes several sources of variation; for example, variation between experiments, between lines, and within lines. The importance of each source depends on the goals of the experimenter, and its impact depends on the design used. With unlimited resources, replication of the complete experiment (including generation of the mapping population) at different times and in different places surely decreases all sources of variation. Whenever resources of fixed, however, it makes sense to choose a design that minimizes the sources of variation that have most impact on the experimental goals. For example, if

one would like to conclude that QTL are repeatable across several independent experiments (perhaps across seasons or locations), then several studies must be performed. If the fine mapping of QTL is the primary goal, then number of recombinant lines that are used should be maximized and experimental variance minimized, often in a single large experiment. In this case, the experiment has been performed only once and, with the resulting data alone, one cannot conclude that the QTL are repeatable. QTL must be confirmed in near isogenic lines (NILs) or heterogeneous inbred families (HIFs) [20,21]. NILs contain a small introgressed fragment in an isogenic background, whereas HIFs are derived from a single recombinant inbred line that segregates a single QTL region in an inbred background that is a mixture of the two parents.

Once QTL are confirmed, they can be characterized further in several environmental conditions and/or genetic backgrounds. The NIL or HIF is also the starting material for the fine-mapping and cloning of the QTL. The availability of plentiful polymorphisms is a boon for fine-mapping because marker discovery is often rate-limiting at this stage. In some cases, the QTL can be mapped directly to the gene [22,23]. Usually, the selection of candidate genes can begin once QTL have been localized to a relatively narrow region (3 cM or less).

Candidate genes

When a full-genome sequence is available, perusing the annotation can often suggest genes in the QTL interval for further study. Predicted functions and gene ontologies help to guide the selection of candidate genes. The process of selecting candidate genes relies on a wealth of information gained through traditional genetics and molecular approaches. Keeping gene annotation up to date with current publications is an important task. Recently, there have been some successful examples of the use of the candidate gene approach to identify QTL genes. The gene encoding the CRYPTOCHROME2 photoreceptor was shown to be responsible for the phenotypic variation associated with a flowering-time QTL [24]. In rice, three QTL have been identified as candidate genes [25–27] whose function was known from studies of *Arabidopsis*.

Once a candidate gene is selected, the first follow-up experiment is usually to sequence the gene in the two parental lines and to look for variation that is predicted to have a functional consequence. As few QTL have been cloned, it is hard to make generalizations about what kind of changes will have phenotypic consequences, but certainly nonsense polymorphisms and deletion polymorphisms make the candidate gene more likely. Amino-acid changes [24,28], as well as expression level changes, may also be important in providing functional variation [29–31]. Several functional alleles have been identified at some QTL loci [32–36]. In such cases, the previous identification of high-density polymorphisms allows the interval to be screened for changes that might have functional consequences [37]. In this regard, genomic DNA hybridization to arrays can reveal changes and potential deletions in genes that make excellent candidates [10••]. A new flowering-time QTL has been identified by this approach in our group (J Werner *et al.*, unpublished).

Gene expression studies in which the NIL QTL is compared to that from the parental line (or an alternative QTL allele from another HIF) for differences in gene expression can also be used to identify candidate genes. Several replicate lines are used to control for biological variance and potential maternal effects. The conditions and tissue selected for the gene expression study must be chosen on the basis of the phenotype of the QTL. Experiments that look at differences in gene expression under several conditions will be more powerful. A set of conditions in which the QTL has no effect provides a control for changes that are unrelated to the phenotype; however, changes in constitutive gene expression may also suggest QTL candidate genes. Gene expression studies also characterize the downstream transcriptional response of the QTL. Thus, genes with expression-level differences that map to the QTL are candidate genes, whereas genes that map to other locations are part of the molecular phenotype caused by the QTL. An alternative experimental design involves the use of lines from the mapping population that have extreme phenotypes. Replicate pools of extreme lines can be profiled independently, so that differences in gene expression will be specific to the phenotype and genotype that separates the pools. This strategy was recently used to identify candidate genes for drought response QTL in rice (S Hazen, personal communication). Large-scale studies are underway to map QTL for gene-expression differences (eQTL) by individually profiling lines from a mapping population. Often, the eQTL map to the gene itself, indicating that *cis* changes are responsible for the different levels of expression; however, the presence of groups of genes that are coordinately regulated by a single unlinked QTL suggests that *trans*-acting factors are controlling expression [38••,39•].

QTL gene confirmation

Once candidate genes have been identified, they need to be tested functionally. A first test of gene function is to identify a null mutation. In *Arabidopsis*, thanks to several massive functional-genomics projects, we have a near-saturating collection of sequence-indexed T-DNA disruption mutants (http://signal.salk.edu; [40•,41••]). More than one null allele can often be identified for most candidate genes within the QTL interval, and the quantitative phenotypes of these alleles can subsequently be measured. The ultimate step in QTL confirmation is to reintroduce alternate alleles (using transgenic techniques)

into reciprocal QTL lines or null mutant backgrounds to show that each allele has a significantly different effect on the phenotype. To date, this has been done for at least two plant QTL [24,42]. Another elegant way to confirm a QTL gene is to use gene replacement, which has been demonstrated successfully in rice [43]. Gene replacement can be used to specifically substitute alleles at the QTL locus while maintaining the correct genomic context, as was performed recently in *Drosophila* [44•].

Conclusions

Several genomics tools are available in *Arabidopsis* that facilitate QTL mapping and cloning (Figure 1). What can be done if an organism does not have a complete genomic sequence? One approach is to use synteny with a relative that has a sequenced genome to identify candidate genes in the region of the QTL [45]. If ESTs are available, they should be mapped so that they can also serve as candidate genes should they fall within the QTL region. One approach to map ESTs quickly is to use oligonucleotide arrays designed from ESTs. Customized arrays can now be designed with no up-front costs [46•]. The hybridization of genomic DNA from parental lines will identify polymorphisms in 25mer features on the array that correspond to these ESTs. Oligonucleotide arrays can then be used to create a high-density genetic map by genotyping a mapping population. This high-density genetic map will resolve the location of many of the ESTs that can serve as candidate genes. Furthermore, these arrays have a dual purpose as they can also be used in expression studies to identify candidate genes. It should be possible to use arrays designed for closely related species for both of these purposes. Spotted arrays may also be effective for mapping ESTs if polymorphisms can be detected. Deletion lines [47] can be used to map ESTs on any array format.

The next five years should see a burst in the number of QTL cloned, thanks to advances in genomics. These QTL will reveal new genes and new alleles of known genes that have evolved in particular genetic backgrounds under specific environmental pressures.

Acknowledgements

We thank Sam Hazen, Charles Berry, Todd Michael, Chris Schwartz and Olivier Loudet for comments on the manuscript. JOB is supported by a fellowship from the Helen Hay Whitney Foundation. Work in our laboratory is supported by the National Institutes of Health. JC is an investigator of the Howard Hughes Medical Institute.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Lander ES, Botstein D: **Strategies for studying heterogeneous genetic traits in humans by using a linkage map of restriction fragment length polymorphisms**. *Proc Natl Acad Sci USA* 1986, **83**:7353-7357.

2. Mauricio R: **Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology**. *Nat Rev Genet* 2001, **2**:370-381.

3. Wang RL, Stec A, Hey J, Lukens L, Doebley J: **The limits of selection during maize domestication**. *Nature* 1999, **398**:236-239.

4. • Jander G, Norris SR, Rounsley SD, Bush DF, Levin IM, Last RL: ***Arabidopsis* map-based cloning in the post-genome era**. *Plant Physiol* 2002, **129**:440-450.
More than 50 000 candidate SNP and indel polymorphisms are revealed by a 2–3× shotgun coverage of the *A. thaliana* accession Landsberg erecta.

5. • Schmid KJ, Sorensen TR, Stracke R, Torjek O, Altmann T, Mitchell-Olds T, Weisshaar B: **Large-scale identification and analysis of genome-wide single-nucleotide polymorphisms for mapping in *Arabidopsis thaliana***. *Genome Res* 2003, **13**:1250-1257.
The authors release more than 8600 polymorphisms from up to 12 accessions that can be searched at http://www.mpiz-koeln.mpg.de/masc/.

6. Peters JL, Constandt H, Neyt P, Cnops G, Zethof J, Zabeau M, Gerats T: **A physical amplified fragment-length polymorphism map of *Arabidopsis***. *Plant Physiol* 2001, **127**:1579-1589.

7. Steemers FJ, Ferguson JA, Walt DR: **Screening unlabeled DNA targets with randomly ordered fiber-optic gene arrays**. *Nat Biotechnol* 2000, **18**:91-94.

8. Shchepinov MS, Denissenko MF, Smylie KJ, Worl RJ, Leppin AL, Cantor CR, Rodi CP: **Matrix-induced fragmentation of P3′-N5′ phosphoramidate-containing DNA: high-throughput MALDI-TOF analysis of genomic sequence polymorphisms**. *Nucleic Acids Res* 2001, **29**:3864-3872.

9. Alderborn A, Kristofferson A, Hammerling U: **Determination of single-nucleotide polymorphisms by real-time pyrophosphate DNA sequencing**. *Genome Res* 2000, **10**:1249-1258.

10. •• Borevitz JO, Liang D, Plouffe D, Chang HS, Zhu T, Weigel D, Berry CC, Winzeler E, Chory J: **Large-scale identification of single-feature polymorphisms in complex genomes**. *Genome Res* 2003, **13**:513-523.
The authors describe a method for identifying and genotyping thousands of markers using arrays that are designed for expression analysis. Potential deletions can be predicted and mutations can be quickly mapped with bulk-segregant analysis and array genotyping.

11. Winzeler EA, Richards DR, Conway AR, Goldstein AL, Kalman S, McCullough MJ, McCusker JH, Stevens DA, Wodicka L, Lockhart DJ *et al.*: **Direct allelic variation scanning of the yeast genome**. *Science* 1998, **281**:1194-1197.

12. Stickney HL, Schmutz J, Woods IG, Holtzer CC, Dickson MC, Kelly PD, Myers RM, Talbot WS: **Rapid mapping of zebrafish mutations with SNPs and oligonucleotide microarrays**. *Genome Res* 2002, **12**:1929-1934.

13. Mackay TF: **The genetic architecture of quantitative traits**. *Annu Rev Genet* 2001, **35**:303-339.

14. Darvasi A: ***In silico* mapping of mouse quantitative trait loci**. *Science* 2001, **294**:2423.

15. Nordborg M, Borevitz JO, Bergelson J, Berry CC, Chory J, Hagenblad J, Kreitman M, Maloof JN, Noyes T, Oefner PJ *et al.*: **The extent of linkage disequilibrium in *Arabidopsis thaliana***. *Nat Genet* 2002, **30**:190-193.

16. Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler ES IV: ***Dwarf8* polymorphisms associate with variation in flowering time**. *Nat Genet* 2001, **28**:286-289.

17. Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler ES IV: **Structure of linkage disequilibrium and phenotypic associations in the maize genome**. *Proc Natl Acad Sci USA* 2001, **98**:11479-11484.

18. • Flint Garcia SA, Thornsberry JM, Buckler ES IV: **Structure of linkage disequilibrium in plants**. *Annu Rev Plant Biol* 2003, **54**:357-374.
A current and comprehensive review of LD mapping approaches in plants.

19. Whitt SR, Wilson LM, Tenaillon MI, Gaut BS, Buckler ES IV: **Genetic diversity and selection in the maize starch pathway**. *Proc Natl Acad Sci USA* 2002, **99**:12959-12962.

20. Loudet O, Chaillou S, Camilleri C, Bouchez D, Daniel-Vedele F: **Bay-0 × Shahdara recombinant inbred line population: a powerful tool for the genetic dissection of complex traits**. *Theor Appl Genet* 2002, **104**:1173-1184.

21. Alonso-Blanco C, Koornneef M: **Naturally occurring variation in *Arabidopsis*: an underexploited resource for plant genetics**. *Trends Plant Sci* 2000, **5**:22-29.

22. Fridman E, Pleban T, Zamir D: **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 USA* 2000, **97**:4718-4723.

23. Kroymann J, Textor S, Tokuhisa JG, Falk KL, Bartram S, Gershenzon J, Mitchell-Olds T: **A gene controlling variation in *Arabidopsis* glucosinolate composition is part of the methionine chain elongation pathway**. *Plant Physiol* 2001, **127**:1077-1088.

24. El-Din El-Assal S, Alonso-Blanco C, Peeters AJ, Raz V, Koornneef M: **A QTL for flowering time in *Arabidopsis* reveals a novel allele of *CRY2***. *Nat Genet* 2001, **29**:435-440.

25. Kojima S, Takahashi Y, Kobayashi Y, Monna L, Sasaki T, Araki T, Yano M: ***Hd3a*, a rice ortholog of the *Arabidopsis FT* gene, promotes transition to flowering downstream of *Hd1* under short-day conditions**. *Plant Cell Physiol* 2002, **43**:1096-1105.

26. Takahashi Y, Shomura A, Sasaki T, Yano M: ***Hd6*, a rice quantitative trait locus involved in photoperiod sensitivity, encodes the alpha subunit of protein kinase CK2**. *Proc Natl Acad Sci USA* 2001, **98**:7922-7927.

27. Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y *et al.*: ***Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS***. *Plant Cell* 2000, **12**:2473-2484.

28. Maloof JN, Borevitz JO, Dabi T, Lutes J, Nehring RB, Redfern JL, Trainer GT, Wilson JM, Asami T, Berry CC *et al.*: **Natural variation in light sensitivity of *Arabidopsis***. *Nat Genet* 2001, **29**:441-446.

29. Kliebenstein DJ, Lambrix VM, Reichelt M, Gershenzon J, Mitchell-Olds T: **Gene duplication in the diversification of secondary metabolism: tandem 2-oxoglutarate-dependent dioxygenases control glucosinolate biosynthesis in *Arabidopsis***. *Plant Cell* 2001, **13**:681-693.

30. Doebley J, Stec A, Hubbard L: **The evolution of apical dominance in maize**. *Nature* 1997, **386**:485-488.

31. Liu J, Cong B, Tanksley SD: **Generation and analysis of an artificial gene dosage series in tomato to study the mechanisms by which the cloned quantitative trait locus *fw2.2* controls fruit size**. *Plant Physiol* 2003, **132**:292-299.

32. Sheldon CC, Burn JE, Perez PP, Metzger J, Edwards JA, Peacock WJ, Dennis ES: **The *FLF* MADS box gene: a repressor of flowering in *Arabidopsis* regulated by vernalization and methylation**. *Plant Cell* 1999, **11**:445-458.

33. Michaels SD, Amasino RM: ***FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering**. *Plant Cell* 1999, **11**:949-956.

34. Le Corre V, Roux F, Reboud X: **DNA polymorphism at the *FRIGIDA* gene in *Arabidopsis thaliana*: extensive nonsynonymous variation is consistent with local selection for flowering time**. *Mol Biol Evol* 2002, **19**:1261-1271.

35. Johanson U, West J, Lister C, Michaels S, Amasino R, Dean C: **Molecular analysis of FRIGIDA, a major determinant of natural variation in *Arabidopsis* flowering time**. *Science* 2000, **290**:344-347.

36. Gazzani S, Gendall AR, Lister C, Dean C: **Analysis of the molecular basis of flowering time variation in *Arabidopsis* accessions**. *Plant Physiol* 2003, **132**:1107-1114.

37. Andersen JR, Lubberstedt T: **Functional markers in plants**. *Trends Plant Sci* 2003, **8**:554-560.

38. •• Schadt EE, Monks SA, Drake TA, Lusis AJ, Che N, Colinayo V, Ruff TG, Milligan SB, Lamb JR, Cavet G *et al.*: **Genetics of gene expression surveyed in maize, mouse and man**. *Nature* 2003, **422**:297-302.

This paper describes expression QTL (eQTL) mapping in three species. Variation in gene expression can be attributable to both *cis- and trans*-acting factors.

39. • Brem RB, Yvert G, Clinton R, Kruglyak L: **Genetic dissection of transcriptional regulation in budding yeast**. *Science* 2002, **296**:752-755.

The authors demonstrate the genetic mapping of expression-level polymorphisms for the first time and identify both co-regulated genes that are controlled by *trans*-acting factors and unique loci that are controlled *in cis*.

40. • Sessions A, Burke E, Presting G, Aux G, McElver J, Patton D, Dietrich B, Ho P, Bacwaden J, Ko C *et al.*: **A high-throughput *Arabidopsis* reverse genetics system**. *Plant Cell* 2002, **14**:2985-2994.

The authors describe a sequence-indexed collection of approximately 53 000 T-DNA lines that have recently been made publicly available.

41. •• Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R *et al.*: **Genome-wide insertional mutagenesis of *Arabidopsis thaliana***. *Science* 2003, **301**:653-657.

This paper describes the creation of a sequence-indexed collection of knock-out lines for nearly all *Arabidopsis* genes (http://signal.salk.edu). At present, more than 140 000 lines have been sequenced.

42. Frary A, Nesbitt TC, Grandillo S, Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB, Tanksley SD: ***fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size**. *Science* 2000, **289**:85-88.

43. Terada R, Urawa H, Inagaki Y, Tsugane K, Iida S: **Efficient gene targeting by homologous recombination in rice**. *Nat Biotechnol* 2002, **20**:1030-1034.

44. • Greenberg AJ, Moran JR, Coyne JA, Wu C-I: **Ecological adaptation during incipient speciation revealed by precise gene replacement**. *Science* 2003, **302**:1754-1757.

Gene replacement is used to confirm the phenotypic effect of allelic differences at a locus that is involved in speciation.

45. Hazen SP, Hawley RM, Davis GL, Henrissat B, Walton JD: **Quantitative trait loci and comparative genomics of cereal cell wall composition**. *Plant Physiol* 2003, **132**:263-271.

46. • Nuwaysir EF, Huang W, Albert TJ, Singh J, Nuwaysir K, Pitas A, Richmond T, Gorski T, Berg JP, Ballin J *et al.*: **Gene expression analysis using oligonucleotide arrays produced by maskless photolithography**. *Genome Res* 2002, **12**:1749-1755.

The authors show that moveable mirrors can be used to synthesize custom oligo-nucleotide arrays *in situ*, allowing greater flexibility in extension to non-model organisms.

47. Qi L, Echalier B, Friebe B, Gill BS: **Molecular characterization of a set of wheat deletion stocks for use in chromosome bin mapping of ESTs**. *Funct Integr Genomics* 2003, **3**:39-55.