REVIEW

Molecular Dissection of Quantitative Traits: Progress and Prospects

Andrew H. Paterson

Department of Soil and Crop Science, Texas A&M University, College Station, Texas 77843-2474

The “molecular dissection” of quantitative traits began before the demonstration that DNA is the hereditary molecule. Opponents of the Mendelian genetic (particulate) theory argued that many traits showed “blending inheritance,” with progeny intermediate between parents. This precipitated one of the great debates in the history of genetics, which was eventually resolved by the realization that blending inheritance might be accounted for by the independent transmission of many different Mendelian factors, together with the modifying effects of environment.

By early in the twentieth century, associations of genetic markers with differences in quantitative phenotypes had already been noted. For example, Sax (1923) noted association of differences in bean seed weight with seed-coat pigmentation, concluding that “Factor differences for seed weight are . . . linked with . . . factors which determine the color of the pigment. Size differences may be effected by the independent action of . . . factors in different linkage groups. These factors, when combined, have a cumulative effect. The size factors in different chromosomes may not be equal in their effect.”

The principles outlined by these pioneering studies were widely employed, both in model systems and domesticated species. However, until recently most maps contained too few genetic markers to permit either identification of suites of quantitative trait loci (QTLs) accounting for the bulk of variation in a phenotype, or precise mapping of QTLs to specific chromosomal locations. As stated by Thoday (1961), “The main practical limitation of the technique seems to be the availability of suitable markers, and the time that can be given to the considerable work involved.”

A growing body of molecular tools, first from protein polymorphisms (Markert and Moller 1959), and, more recently, from DNA polymorphisms, has contributed to the development of “complete” genetic maps, encompassing all regions of all chromosomes in a plant or animal species. Complete genetic maps permit comprehensive analysis of the genome of an organism, revealing the locations of QTLs influencing virtually any characteristic that can be measured.

The development of DNA-based genetic mapping technology was propelled by the search for genes responsible for human genetic diseases and other traits (Botstein et al. 1980). QTLs influencing medically important phenotypes such as high blood pressure (Rapp et al. 1989) and hypertension (Jacob et al. 1991) are being identified in mammalian models. Recent results have illustrated the usefulness of genome mapping in dissecting complex behavioral characteristics (Plomin et al. 1994), such as avoidance, exploration (Neiderheiser et al. 1992), and substance abuse (Crabbe et al. 1994; Quock et al. 1994) in rodents, and reading disability in humans (Cardon et al. 1994). In one recent case, the effectiveness of a human disease agent (malaria) at parasitizing its vector (mosquito) has been dissected into several QTLs (Severson et al. 1995). Finally, many complex traits mapped in domestic animals, such as obesity (Andersson et al. 1994; Pelleymounter et al. 1995), lactation (Georges et al. 1995), and others (cf. Cockett et al. 1994; Georges et al. 1994) may be relevant to human phenotypes. An ever-clearer picture of the correspondence between the chromosomes of humans and other mammals helps in the application of such comparative information (cf. O’Brien et al. 1993).

Plant breeding has provided a fertile field for molecular dissection of quantitative traits. Charles Darwin recognized that plant breeding programs were very useful for the study of heredity. The realization that crop improvement could benefit from DNA marker-assisted selection (cf. Paterson et al. 1991b) has led to rapid growth in molecular mapping of agriculturally important crops. Genetic mapping of traits associated with many aspects of crop productivity is published or
in progress, including basic growth patterns such as plant height (cf. Koester et al. 1993; Lin et al.
1995; M. Pereira, M. Lee, and P. Rayapati, un-publ.), tillering, rhizomatousness (Paterson et al.
1995a), flowering time (cf. Koester et al. 1993; Kowalski et al. 1994; Li et al. 1995a; Lin et al.
1995), and other morphological variants (Kennard et al. 1994); yield components such as the
size, number, and harvestability of seed (Stuber et al. 1987, 1992; Abler et al. 1991; Fatokun et al.
1992; Doebley et al. 1994; Schoi et al. 1994; Paterson et al. 1995a,b), biomass, and/or growth
rates (DeVicente and Tanksley 1993; Bradshaw and Stettler 1995); quality parameters such as
composition of fruit or seed (Paterson et al. 1988, 1990, 1991a; Weller et al. 1988; DeVicente and
Tanksley 1993; Teutonico and Osborn 1994), shape of tubers (Van Eck et al. 1994), or specific
gravity of wood (Groover et al. 1994); and the impact of adverse factors such as diseases (Bubeck
et al. 1993; Leonards-Schippers et al. 1994; Wang et al. 1994; Jung et al. 1995; Li et al. 1995b), in-
sects (Nienhuis et al. 1987; Bonierbale et al. 1994) and abiotic factors (Martin et al. 1989; Reiter et
al. 1991), and the evolution of novel organs (Doebley et al. 1990). This partial list illustrates
the breadth of topics in crop improvement in which QTL mapping is affording new insights.

In addition to medicine and agriculture, QTLs playing critical roles in evolution have been
characterized. Bradshaw et al. (1995) recently mapped QTLs controlling floral traits that de-
termine pollinator preference in two Mimulus species. These species are sympatric (grow in the
same places) and can produce fertile hybrid progeny—but do not do so in nature, because one
species is strictly bumblebee pollinated, and the other is strictly hummingbird pollinated. A few
QTLs with large effects determine the floral characteristics (flower color, nectar volume and con-
centration, and stamen and pistil length) that in-
fluence pollinator preference—these mutations
(QTLs) may represent the initial steps in repro-
ductive isolation of these two species.

Analysis and Interpretation of QTL Mapping Experiments

Algorithms have been developed for QTL map-
ing in a wide range of pedigrees and experimen-
tal designs, including F2, backcross, recombinant
inbred, and many other designs (cf. Weller 1986;
Lander and Botstein 1989; Knapp et al. 1991; Luo
Mackinnon and Georges 1992; Darvasi et al.
1993; Moreo-Gonzalez 1993; Rodolphe and Le-
fort 1993; Zeng 1993, 1994; Cardon and Fulker
1994; Eaves 1994; Haley et al. 1994; Jiang and
Zeng 1995). All share the basic principle of test-
ing correlation between marker genotypes and
quantitative phenotypes. As a result of the avail-
ability of complete genetic maps, current analyti-
cal methods are able to use information from
multiple markers that flank a QTL, in contrast to
earlier methods that were limited to information
from single markers at unknown distance and di-
rection from the QTL. This permits more accurate
estimates of location and phenotypic effect, of
individual QTLs. Although many procedures to
date have been parametric, approaches to han-
dling nonQTLs are beginning to emerge
(Kruglyak and Lander 1995).

Perhaps the single most important consider-
ation in analysis and interpretation of QTL data
is the threshold employed for inferring that a
QTL is statistically significant. Because QTL mapping involves many analyses of independent (un-
linked) genetic markers throughout a genome,
there are many opportunities for false-positive
results. Stringent significance thresholds must be
employed to avoid these. Nominal significance
criteria of 99.8% or more for any single QTL are
usually necessary to assure an “experiment-wide”
confidence level of 95% for all QTLs reported
across a genome. Appropriate criteria are often
described in detail when an analytical approach
is developed (cf. Lander and Botstein 1989). Al-
ternatively, methods for empirical calculation of
criteria appropriate to particular data sets have
been described (Churchill and Doerge 1994; Re-
bai et al. 1994, 1995). As new opportunities for
“comparative analysis” of previously published
QTLs emerge (cf. Lin et al. 1995; Paterson et al.
1995b), it becomes ever more important that
published QTLs satisfy statistical criteria that
minimize the likelihood of false-positive results.

Genetic Basis of Quantitative Variation

QTL mapping is useful for investigating specific
properties of individual genes contributing to
quantitative traits. In contrast, classical quantita-
tive genetics describes the aggregate behavior of
suites of genes influencing a trait. The aggregate
descriptions of quantitative genetics are very use-
ful for guiding manipulation of plant and animal
gene pools by breeders; however, an understand-
ing of quantitative inheritance at the molecular
level requires detailed descriptions of individual
genes, and this is made possible by QTL mapping.

The reconciliation of results from QTL map-
ing with expectations from classical theory is
progressing but remains incomplete. Basic prop-
erties of individual QTLs such as additivity and
dominance are readily tested, and a wide range of
different modes of gene-action are evident for
different QTLs (cf. Paterson et al. 1991a). Many
basic genetic phenomena are readily reconciled
with results from QTL mapping, such as trans-
gression (progeny that exceeds parental genetic
potential; Paterson et al. 1988; deVicente and
Tanksley 1993) and heterosis (hybrid vigor; Stu-
ber et al. 1992; Xiao et al. 1995). However, two
particular areas continue to be controversial: the
number of genes with pronounced effects on
quantitative traits and the importance of interlo-
cus interactions ("epistasis").

Gene Numbers and Phenotypic Effects
Geneticists have long debated the degree of com-
plexity of quantitative traits (cf. Dove 1993).
Many theories, ranging from "virtually infinite
numbers of genes with tiny effects", to "few
genes with large effects" have been proposed,
championed, questioned, revised, rejected, and
reincarnated. Geneticists realized that some as-
sumptions invoked to simplify quantitative mod-
els, such as equality of gene effects and strict ad-
dditivity of gene action, were unlikely to describe
individual quantitative trait loci precisely. Such
assumptions were tolerated so long as models
were reasonably accurate in predicting the be-
havior of study populations. It was no particular sur-
prise that QTL mapping showed such assump-
tions to be incorrect. However, it has remained
controversial whether the results of QTL map-
ing experiments reflect the true complexity of
quantitative inheritance or simply detect a subset of
(relatively large) gene effects.

Figure 1 shows histograms of the fractional
phenotypic variance explained by QTLs affecting
several quantitative traits, mapped in sorghum
and rice. Similar data have been published previ-
ously for tomato (Paterson et al. 1991a). In each
case, a relatively small number of genes accounts
for very large portions of phenotypic variance,
with increasing numbers of genes accounting for
progressively smaller portions of variance, until
the significance threshold is reached. Curiously,
the number of QTLs observed in "first-generation
mapping" (e.g., of a BC1 or F2 population; cf.
Paterson et al. 1988) is often similar to the "ef-
fective number of factors" predicted by classical
models. However, if genes explaining large por-
tions of phenotypic variance are rendered ho-
mozygous, the significance threshold is reduced,
and yet additional genes explaining even smaller
portions of phenotypic variance are revealed
(Paterson et al. 1990).

Although the "end" of this stepwise process of
revealing genes of smaller and smaller effect is
not clear, a large and growing body of data con-
tinues to suggest that relatively few genes ac-
count for the bulk of variation in many popula-
tions, with ever-larger numbers of genes contrib-
uting ever-smaller portions of variance (cf. Lande
and Thompson 1990). This result will be sup-
ported further in a discussion of comparative
analysis, with the finding that genes accounting
for the bulk of variation in a trait appear to cor-
respond in different reproductively isolated spe-
cies.

Epistasis
Almost universally, the collective activities of
mapped QTLs explain only a portion of the phe-
notypic difference between parents, even with nearly complete genome coverage by DNA markers. Classical evidence has strongly suggested the importance of epistasis, or nonlinear interactions between unlinked genetic loci, in quantitative inheritance (cf. Spickett and Thoday 1966; Falconer 1981; Mather and Jinks 1982; Pooni et al. 1987; Allard 1988).

Until recently, QTL mapping experiments have provided very little evidence in support of the importance of epistasis, with nonlinear interactions among DNA marker loci reaching statistical significance at approximately the frequency that would be expected to occur by chance (cf. Edwards et al. 1987; Paterson et al. 1988, 1991a). Hints of epistasis among QTLs have come from the demonstration of “genetic background effects” on quantitative traits in Drosophila (Spassky et al. 1965), rice (Kinoshita et al. 1982, Sato and Sakamoto 1983), and tomato (Tanksley and Hewitt 1988) and from the discovery of occasional loci reported to show interaction with multiple unlinked sites in a genome (cf. Paterson et al. 1988).

Modified experimental designs may reconcile QTL mapping results with the importance attributed to epistasis in classical studies. Doebley et al. (1995) developed genetic stocks differing by two QTLs suspected to interact epistatically, but otherwise uniform in genetic background, and found strong evidence for epistasis between the loci. Lark et al. (1995) used recombinant inbred lines to reduce the complexity of interactions and replicate phenotypic measurements, and found evidence of epistasis between QTLs in genetic control of several agronomic traits. Each of these results suggests that the absence of epistasis in previous QTL mapping studies may have been the result of insufficient replication and/or insufficient statistical resolution to detect interactions, in the presence of many QTLs with large main effects. Although the effects of some QTLs appear independent of interacting loci (cf. Paterson et al. 1990), in at least some cases it is becoming clear that “the whole” is, indeed, greater than the sum of the parts. Improved documentation of epistasis may account for a portion of the “genetic difference between parents” that was unexplained previously by QTL mapping.

It is noteworthy that analysis of epistasis is even more subject to the problem of false-positive results than is analysis of individual QTLs (as discussed above). An important need in QTL research is a generally applicable set of criteria for inferring statistical significance of interactions between genetic loci, which controls “experiment-wise” error rates.

Comparative Analysis of QTLs

“Comparative mapping,” the alignment of the chromosomes of related taxa based on genetic mapping of common DNA markers, affords many benefits to genome analysis. Many DNA probes can be cross-utilized among different species in the same taxonomic family, increasing the number and map density of genetic markers available for many genera simultaneously. In one well-studied plant family, the Poaceae, extensive conservation of gene repertoire and order (cf. Hulbert et al. 1990; Ahn and Tanksley 1993; Ahn et al. 1993a,b; Kurata et al. 1994) has led to the suggestion that the cultivated cereals (diverse genera within the family Poaceae) might be treated as essentially a “single genetic system” (Bennetzen and Freeling 1993; Helentjaris 1993).

Comparative mapping is also useful in molecular dissection of quantitative traits (Fig. 2). A close correspondence among QTLs affecting complex traits such as seed size, as well as traits of varying complexity in different taxa such as disarticulation (“shattering”) of the inflorescence and day-neutral flowering, has been shown for sorghum, sugarcane, maize, wheat, barley, and rice (Paterson et al. 1995b). Correspondence among an unexpectedly high proportion of genes affecting height and flowering of maize, sorghum, and other grasses has also been reported (Lin et al. 1995a).
OIL MAPPING: PROGRESS AND PROSPECTS

Maize Chromosome 1

Maize Chromosome 5

Sorghum Linkage Group C

Maize Chromosome 9

Figure 2 (See facing page for legend.)
Correspondence of QTLs in different species of the plant genera *Lycopersicon* (Paterson et al. 1991a) and *Vigna* (Fatokun et al. 1992) had been reported previously; however, this correspondence spans rather short periods of genetic divergence, and the relative promiscuity of plant "species" makes it difficult to preclude the possibility of recent gene flow.

Correspondence of QTLs on duplicated chromosome segments within a particular species has also been suggested, indicating that chromosome duplication may contribute to polygenic inheritance. Specifically, pairs of loci affecting shattering of the maize inflorescence (see Fig. 2; Doebley et al. 1990; Paterson et al. 1995b), maize height and flowering, and sorghum height (Lin et al. 1995) fall on corresponding duplicated chromosome segments. Chromosome duplication in these taxa preceded domestication by millions of years. Domestication, in the past 10,000 years, most likely imparted strong selective advantages to reduced shattering, reduced height, and early flowering. Consequently, if duplicated genes retained common functions, it is intuitive that QTLs might be found on homologous (duplicated) chromosomal sites (Paterson et al. 1995b). Ongoing research in these and other polyploids is likely to reveal additional cases of putatively duplicated QTLs.

**Implications of QTL Correspondence**

The suggestion that mutations in corresponding genes may account for phenotypic variation in taxa that have been reproductively isolated for millions of years has many implications. Perhaps first and foremost, QTL analysis in one taxon may predict results in other taxa. Such predictive value would make QTL mapping results more broadly applicable than was previously envisioned, enabling research on facile systems to be extrapolated to more difficult ones.

Correspondence of QTLs across diverse taxa also provides strong empirical support for the use of model systems in research on complex phenotypes. For example, the ease of genetic analysis in rodents and domesticated mammals has permitted mapping of genes associated with diabetes, hypertension, obesity, alcohol/drug addiction, and other medically important phenotypes. The difficulties associated with mapping complex traits in humans may be partly overcome by the possibility of cloning QTLs from mouse or other mammals that also account for phenotypic variation in humans. In a similar manner, crop plants that grow particular organs of extraordinary size, such as the enlarged root of turnip, inflorescence ("curd") of cauliflower, or fruit of tomato, might be used to isolate genes important to particular aspects of plant growth and development.

Finally, correspondence among QTLs in different taxa strongly supports the hypothesis that a relatively small number of mutations (genes) accounts for a large portion of phenotypic variation in many populations. If the possibilities for such mutations were infinite and all equally probable, such correspondence seems unlikely. In contrast, if a few genes play disproportionately large roles in genetic control of a trait, then mutations in (one of) these genes are more likely to have effects sufficiently large to drive the mutant allele to fixation.

It remains unclear whether the correspondence among QTLs that is found in interspecific crosses (cf. Paterson et al. 1991a; Fatokun et al. 1992; Lin et al. 1995; M. Pereia, M. Lee, and P. Rayapati, unpubl.) is paralleled by correspondence of QTLs in more closely related genotypes such as elite crop cultivars (cf. Beavis et al. 1991). This may indicate that crop gene pools are homogeneous at mutant alleles with large effects and fixed during the initial stage of crop domestication and that variation within elite gene pools of individual crops is a result of subsequent mutations at a larger number of loci with smaller effects.

**Toward Cloning of QTLs**

Map-based cloning has been used successfully to isolate mutant alleles imparting discrete phenotypes; however, QTLs continue to be refractory to cloning, for several reasons. Individual QTLs exert a relatively small effect on phenotype, which can be obscured by the effects of other genes or by nongenetic factors that are difficult to control. Breeding approaches can be used to create populations segregating for individual QTLs (cf. Paterson et al. 1990; Dorweiler et al. 1993) but are time and labor intensive. The resolution with which QTLs can be mapped is rather coarse; even using "second-generation" experimental designs based on progeny testing, a resolution of ~3 cM is the best achieved to date. In a genome such as that of human or many crop plants, 3 cM represents ~0.1% of the genome, containing perhaps 100 or
more genes. Identification of these genes can be difficult (Gardiner and Mural 1995). Furthermore, complementation testing of so many genes, requiring transformation of the dominant allele into the recessive genotype, is an enormous undertaking, even in taxa that enjoy good transformation systems.

Several recent developments make cloning of QTLs more tractable. First, the technology for cloning and manipulating large segments of intact chromosomal DNA is becoming ever more efficient, facilitating “chromosome walks” (Steinmetz et al. 1982). A decade ago, chromosome walks were based on phage and cosmid vectors with maximal carrying capacity of 15,000–30,000 nucleotides of exogenous DNA. “Artificial chromosomes” carrying 10 times this amount of DNA can now be developed routinely and are publicly available for several organisms. Initially, artificial chromosomes containing exogenous DNA were propagated in yeast (Burke et al. 1987); however, bacterial systems (Shizuya et al. 1992) are quickly gaining acceptance. Improved vectors with new combinations of selectable markers and systems that facilitate transformation of megabase DNA segments (cf. Pachnis et al. 1990; Pavan et al. 1990; Eliceriri et al. 1991; Strauss and Jaenisch 1992) will accelerate the search for QTLs, especially in combination with site-specific recombination systems (cf. Sauer and Henderson 1988).

There is a large and growing number of approaches for identifying candidate genes that might be QTLs. Direct isolation of genes from megabase DNA clones is becoming more efficient (Lovett et al. 1991; Parimoo et al. 1991; Collins 1992; Morgan et al. 1992). Several PCR-based techniques are now available for identifying genes expressed specifically in particular tissues, developmental stages, or genotypes (Hara et al. 1991; Liang and Pardee 1992; Sun et al. 1992; Lopez-Fernandez and Mazo 1993; Liang et al. 1994). Moreover, random sequencing of cDNAs is adding to a rapidly growing repertoire of plant genes that are readily accessible (cf. Hofte et al. 1993; Kurata et al. 1994; Newman et al. 1994; Sasaki et al. 1994). By combining such information with genetic mapping to test which candidates are located in the target chromosomal region, isolation of candidate genes might be accelerated.

The identification of multiple independent mutant alleles at QTLs would permit testing of correlation between phenotype and mutation in a particular transcript, as has been used in map-based cloning of genes in humans and other organisms. Using comparative information, genotypes of different taxa might be identified that show a common phenotype that can be explained by mutations mapping to corresponding chromosomal locations (cf. Paterson et al. 1995b; Lin et al. 1995). Sequencing of dominant and recessive alleles at “candidate genes” in each taxon might allow a mutation in a particular transcript to be correlated with the mutant phenotype, in much the same manner as has been employed to identify several disease genes in human.

The suggestion that QTLs might be cloned by identifying allelic mutations with discrete effects (Robertson 1985) has also come closer to realization. Comparative analysis of different taxa provides more opportunities to seek such discrete mutations, and several examples have been reported (Beavis et al. 1991; Lin et al. 1995; Paterson et al. 1995b; M. Pereira, M. Lee, and P. Rayapati, unpubl.). Identification of discrete mutations corresponding to QTLs offers a more rapid and efficient means to dissect quantitative traits than do breeding approaches (Paterson et al. 1990; Dorweiler et al. 1993).

Summary

QTL mapping is an increasingly useful approach to the study and manipulation of complex traits important in agriculture, evolution, and medicine. The molecular dissection of quantitative phenotypes, supplementing the principles of classical quantitative genetics, is accelerating progress in the manipulation of plant and animal genomes. A growing appreciation of the similarities among different organisms and the usefulness of comparative genetic information is making genome analysis more efficient, and providing new opportunities for using model systems to overcome the limitations of less-favorable systems.

The expanding repertoire of techniques and information available for studying heredity is removing obstacles to the cloning of QTLs. Although QTL mapping alone is limited to a resolution of 0.1%–1.0% of a genome, use of QTL mapping in conjunction with a search for mapped candidate genes, with emerging technologies for isolation of genes expressed under conditions likely to account for the quantitative phenotype, and with ever more efficient megabase DNA manipulation and characterization bodes...
PATERN

well for the prospect of isolating the genetic determinants of QTLs in the foreseeable future. In the words of Thoday (1961), “An extensive attack on quantitative genetics made from this point of view as well as the biometric approach should be a great help in answering questions concerning the nature of polygenes . . .”

ACKNOWLEDGMENTS

I thank M. Burow, C. Katsar, and T.-H. Lan for critical reading of the manuscript and numerous colleagues and staff for valuable discussions. Aspects of this research were funded by the International Consortium for Sugarcane Biotechnology, U. S. Department of Agriculture (Plant Genome and Biotechnology Risk Assessment Programs), Pioneer HiBred International, Texas Higher Education Coordinating Board, and Texas Agricultural Experiment Station.

REFERENCES


OIL MAPPING: PROGRESS AND PROSPECTS


Jansen, R.C. 1992. A general mixture model for mapping...


QTL MAPPING: PROGRESS AND PROSPECTS


Received September 18, 1995; accepted in revised form October 26, 1995.
Molecular dissection of quantitative traits: progress and prospects.

A H Paterson

Genome Res. 1995 5: 321-333
Access the most recent version at doi:10.1101/gr.5.4.321

References

This article cites 133 articles, 72 of which can be accessed free at:
http://genome.cshlp.org/content/5/4/321.full.html#ref-list-1

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.