Lessons from a search for leaf mutants in *Arabidopsis thaliana*

JOSÉ MANUEL PÉREZ-PÉREZ, HÉCTOR CANDELA, PEDRO ROBLES, VÍCTOR QUESADA, MARÍA ROSA PONCE and JOSÉ LUIS MICOL*

División de Genética and Instituto de Bioingeniería, Universidad Miguel Hernández, Campus de Elche, Elche, Alicante, Spain

**ABSTRACT** Large-scale exploratory approaches to understanding gene function laid the foundations for the “-omics era”. Based on modern technologies for the structural and functional characterization of genomes, these curiosity-driven approaches allow systematic accumulation of vast amounts of data, enabling subsequent hypothesis-driven research. Some years before the dawn of genomics, exploratory approaches were already furthering our understanding of gene function in the form of saturation mutagenesis experiments aimed at the identification of all genes that mutate to a given phenotype. Forward genetic approaches, conducted on experimental organisms such as *Drosophila melanogaster* and *Caenorhabditis elegans*, have led to the isolation of mutants affected in specific developmental processes, whose cellular and molecular characterization has unraveled the underlying genetic mechanisms of animal development. To shed light on the making of plant leaves, in 1993 we initiated an attempt to identify as many viable and fertile mutants with abnormal leaf morphology as possible, using the *Arabidopsis thaliana* model organism. We identified 25 fast-neutron- and 153 ethyl-methane sulfonate-induced mutations, which fell into eight and 94 complementation groups, respectively. We also studied 115 publicly available mutant lines isolated by previous authors, which fell into 37 complementation groups. Although we did not reach saturation of the *Arabidopsis thaliana* genome, the broad spectrum of leaf morphological alterations identified is facilitating the dissection of specific leaf developmental processes. In a complementary approach, we also analyzed leaf architecture in natural accessions and two populations of recombinant inbred lines. Using a high-throughput gene mapping method, we have already cloned 25 of the genes identified by mutation, in some cases in collaboration with other groups. The products of these genes participate in various developmental processes, such as polar cell expansion, transduction of hormonal signals, gene regulation, plastid biogenesis, and chromatin remodeling, among others. The range of phenotypes and processes identified reveal the complexity of leaf ontogeny and will help explain the diversity of leaf morphology in nature.

**KEY WORDS:** *Arabidopsis thaliana*, forward genetics, leaf development, mutant screen

While walking in the Public Gardens of Palermo, it came to me in a flash that in the organ of the plant which we are accustomed to call the leaf lies the true Proteus who can hide or reveal himself in all vegetal forms. From first to last, the plant is nothing but leaf, which is so inseparable from the future germ that one cannot think of one without the other.

Goethe (1816).

Curiosity-driven versus hypothesis-driven research

The advent of high-throughput, cost-effective techniques for the structural and functional characterization of genomes has led to a revolution in biology. In an unprecedented approach, the emphasis of scientific research is not on the generation and experimental testing of hypotheses, but on the systematic gathering of vast amounts of data (i.e., genome sequences) and the generation of large collections of resources (i.e., tagged mutants, gene libraries), often accessible through publicly available databases. These initial observations provide the starting information.
that enables subsequent hypothesis-driven research, which would be impossible in the absence of any prior knowledge of the process under study. Underlining the fact that no guiding hypothesis is required, this type of research is referred to in various ways, including terms such as exploratory, inductive, observational, data-driven, question-driven, or curiosity-driven. Curiosity-driven approaches are sometimes neglected, even though they have the potential to yield substantial, rather than incremental, advances in knowledge simply because they are not constrained by the limits and expectations imposed by current paradigms. As some authors have highlighted (Kell and Oliver, 2004; Franklin, 2005), curiosity-driven research contributes to the advancement of science more than is usually recognized and is complementary in nature to hypothesis-driven research, since both act as iterative partners in the cycle of knowledge.

Some years before the “-omics era”, exploratory approaches were already furthering our understanding of gene function, in the form of large-scale mutagenesis experiments aimed at the identification of mutant alleles for all genes that contribute to a given phenotype (Murray, 2000). The classical example was pioneered by Nüsslein-Volhard and Wieschaus (1980), who devised the methodology for saturation mutagenesis and applied it to the dissection of embryonic development in the model organism *Drosophila melanogaster*, with no assumption other than that development is under genetic control. Our current understanding of many other important biological phenomena derives from applying the same rationale to many experimental systems. More recently, genomics has provided researchers with a wealth of data from high-throughput array-based experiments (Lockhart and Winzeler, 2000), including global analyses of gene expression under multiple experimental conditions, and from novel massively parallel sequencing techniques (Margulies et al., 2005; Rogers and Venter, 2005), which can be used to obtain direct estimates of the abundance of molecules being sequenced. These techniques are expected to assist in the design of hypotheses that will lead to a mechanistic understanding of the regulatory networks that underlie complex processes, such as the development of multicellular organisms. In this context, the generation of collections of mutants induced by random mutagenesis fulfills the dual objective of providing a survey of the range of phenotypes that can arise after perturbation of the wild-type condition and of being an invaluable resource for further hypothesis-driven research.

**Mutant genetic screens in model organisms: a historical perspective**

Multicellular organisms develop from a unicellular zygote by highly regulated processes that result in progressively increased structural complexity. During embryonic development, a hierarchy of gene products participates in cell specification and morphogenetic pathways that cooperatively build the architecture of the adult body. Understanding how genetic information is integrated and used for the elaboration of complex biological patterns is among the main challenges of developmental biology.

Soon after the rediscovery of Mendel’s laws in 1900, pioneering genetic studies conducted on the fruit fly (*Drosophila melanogaster*) in the laboratory of Thomas H. Morgan led to the discovery that genes are physically located on chromosomes and exist in many allelic forms (reviewed by Sturtevant, 1965). Morphological variation was soon recognized as being caused by spontaneous mutations at very low rates, which could be increased experimentally (Müller, 1927). It is generally agreed that such mutations are a powerful tool for unraveling the normal function of individual genes during development.

Forward genetic approaches begin with the induction and isolation of mutants affected in a given process and then infer the biological role of mutated genes based on their associated phenotypes. In this way, high-throughput isolation of morphological mutants and the cloning of the corresponding genes have increased our knowledge of the events that take place during development in organisms such as fungi, nematodes, insects, vertebrates, and plants. One limitation of mutant screens is that mutations that produce subtle phenotypes are likely to remain undetected. A combination of reverse genetics, clonal analysis, and cell biology techniques provides the toolbox to overcome this problem and will help identify all genetically detectable functions at work during the development of complex organisms.

The fruit fly offers many advantages as an experimental organism for developmental genetic studies, including its small size, short life cycle, high fertility, and relatively small genome (Rubin and Lewis, 2000). Because even subtle wing deformations can easily be detected, mutations altering wing shape or wing venation were among the first to be studied at a large scale (Waddington, 1940). Four decades later, embryonic development of the fruit fly was examined with a saturation mutagenesis screen (Nüsslein-Volhard and Wieschaus, 1980; Nüsslein-Volhard et al., 1984). From approximately 27,000 lines, 580 mutants with defects in the larval segmentation pattern were identified, affecting 139 different genes (Table 1). Many of these genes are now known to encode transcription factors necessary for cell fate specification along the anteroposterior and dorsoventral embryonic axes (reviewed by Nüsslein-Volhard, 1991). These studies highlighted for the first time how a complex morphogenetic pattern emerges from a few local signals or morphogens that regulate hierarchies of downstream transcription factors, which interact to delimit cellular territories in the adult body (Nüsslein-Volhard, 1995). More recently, the use of sophisticated imaging techniques to visualize subtle patterning defects has allowed detailed screens aimed at the genetic dissection of other complex processes, such as neuronal morphogenesis (Gao et al., 1999; Table 1).

Despite its late adoption as a model organism (Brenner, 1974), genetic screens performed with the nematode *Caenorhabditis elegans* were also important for the characterization of cell differentiation pathways, such as the differentiation of the vulva from undifferentiated epidermal cells (Ferguson and Horvitz, 1985; Ferguson et al., 1987). In these screens, 95 mutations defining 22 genes were identified and later found to participate in two highly conserved signaling pathways required for cell fate specification (reviewed by Wang and Sternberg, 2001). The possibility of silencing target genes by RNA interference (RNAi) has greatly accelerated the search for loss-of-function phenotypes in *C. elegans* (Kamath et al., 2003). For instance, a genomewide RNAi screen allowed researchers to identify 661 genes required for the first two rounds of cell division in the nematode embryo (Sönntuchsen et al., 2005; Table 1).

Many genes involved in axial specification in *D. melanogaster*
Leaf mutant screens in Arabidopsis and C. elegans have homologs in vertebrates. The zebrafish (Danio rerio) is the only vertebrate systematically used in forward genetic screens (reviewed by Patton and Zon, 2001). Its translucent embryo is ideal for the identification of mutants deficient in genes required for patterning and differentiation from fertilization to late embryogenesis. In two large-scale screens for visible embryonic phenotypes, more than 1,200 mutants were isolated, which defined at least 550 genes (Driever et al., 1996; Haffter et al., 1996; Table 1). Although these screens failed to saturate the genome with embryonic mutations, they allowed the discovery of mutants with defective blood and heart formation, musculature, and pigmentation. Functional redundancy due to duplicated genes is probably the most severe limitation for determining gene function through mutational approaches in zebrafish (Haffter et al., 1996). New morpholino-based gene silencing protocols have allowed researchers to overcome this problem and systematically knock down selected target genes (Sumanas and Larson, 2002).

Since multicellularity evolved independently in plants and animals, they are expected to use different developmental mechanisms. Like in animals, the basic body plan of plants is laid down during embryogenesis, but, unlike in animals, plant morphogenesis is mostly postembryonic. Mutational approaches became routine in the model plant Arabidopsis thaliana in the early 1990s, after initial screens for embryo-lethal mutants (Meinke and Sussex, 1979). The primary reason for studying seedling lethal mutants with altered spatial embryo organization, initiated by Mayer et al. (1991; Table 1), was to learn about the logic of pattern formation by determining how the embryo responds to the removal of a single, genetically defined component. Although many genes that regulate these early patterning events have been identified, some of them involved in cytokinesis and either auxin signaling or auxin transport, the underlying mechanisms remain poorly understood (reviewed by Jenik et al., 2007).

Genetic analysis has furthered our understanding of the mechanisms responsible for the differentiation of specialized cell types within plant tissues. The ontogeny of trichomes, large polarized cells that are regularly spaced on the leaf surface, has been studied genetically in Arabidopsis thaliana (Hülskamp et al., 1994). In total, 71 mutations that affect trichome morphogenesis and spatial patterning were classified based on the phenotypes produced. Complementation tests indicated that these mutations affect 22 loci (Hülskamp et al., 1994; Table 1). Although additional loci involved in trichome development have been discovered since this early work, an analysis of the cellular phenotypes in single and double mutants has allowed researchers to study specific events in trichome development (reviewed by Schellmann and Hülskamp, 2005) and to propose an elegant model for the differentiation of a specialized cell type from a uniform cell layer. The spatial pattern of presumptive trichome cells is first laid down in the developing leaf, and the committed trichome cells then undergo morphogenesis and differentiation through a defined network of genetic interactions.

Homeotic mutants (i.e., mutants with normal organs in positions typically occupied by another organ type) have facilitated the

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Model organism</th>
<th>Screened for</th>
<th>Mutants recovered</th>
<th>Genetic loci involved</th>
<th>Degree of saturation reached</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drosophila melanogaster</td>
<td>Embryo lethality and larval cuticle defects</td>
<td>580</td>
<td>139</td>
<td>98.8</td>
</tr>
<tr>
<td>Caenorhabditis elegans</td>
<td>Altered cell lineages of the vulva</td>
<td>95</td>
<td>22</td>
<td>98.7</td>
</tr>
<tr>
<td>Danio rerio</td>
<td>Abnormal embryonic and early larval phenotypes</td>
<td>331</td>
<td>220</td>
<td>77.8</td>
</tr>
<tr>
<td>Arabidopsis thaliana</td>
<td>Defects in embryonic development</td>
<td>90</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

aAssuming a Poisson distribution of frequencies (equal mutability for all genes). bOnly mutations on genes in the second chromosome were scored. n.a., not applicable.

---

**Fig.1.** Leaf diversity as seen by Carolus Linnaeus in a page of his *Philosophia Botanica* (1751).
identification of organ identity genes responsible for the whorled structure of Arabidopsis thaliana and Antirrhinum majus flowers (reviewed by Meyerowitz et al., 1991). Studies of single and double homeotic mutants in these species led to the ABC model, which states that the identity of each whorl in a wild-type flower is specified by a unique combination of organ identity genes, all of which encode plant transcription factors (reviewed by Coen and Meyerowitz, 1991; Weigel and Meyerowitz, 1994; Ng and Yanofsky, 2000; Coen, 2001).

A search for leaf mutants in Arabidopsis thaliana

The leaf, one of the most conspicuous organs of higher plants, is the main site of primary productivity and a basic anatomical unit. Plant leaves have determinate growth and develop from a small population of founder cells at the flanks of the shoot apical meristem (Byrne, 2005; Tsukaya, 2005; 2006). In addition to their main role in light harvesting, leaves are also important for nutrient storage, defense, and stress responses. Leaves were considered crucial for taxonomic purposes by Linnaeus (1751) (Fig. 1) and seen as the basic units of plant body architecture by Goethe (1816). It is therefore surprising that in 1993, when the laboratory of José Luis Micol was established, a Medline search using the keywords "Arabidopsis and leaf and mutant and development" yielded only 12 articles. Obviously, plant vegetative development was not receiving much attention from the scientific community at that time. We considered answering questions on how many genes control leaf development and how these genes interact essential to understanding how postembryonic lateral organs are built in plants.

To dissect leaf organogenesis, we followed a forward genetic approach using Arabidopsis thaliana as a model organism. We aimed to identify as many viable and fertile mutants with abnormal leaves as possible, in an attempt to isolate mutant alleles of most genes involved in leaf development. The Landsberg erecta (Ler; Fig. 3A) accession was selected as the genetic background for our mutagenesis because it is easily grown at high density in the laboratory, has an upright inflorescence that facilitates crossing and seed harvesting, and because the polymorphism between the Ler and Columbia-0 (Col-0) genomes makes genetic mapping and map-based cloning straightforward. The alkylating agent ethyl-methane sulfonate (EMS), by far the most efficient and commonly used mutagen, was chosen because it induces G-to-T transitions and has no known bias for particular genomic regions (Greene et al., 2003). Based on codon usage in the Arabidopsis thaliana genome, the frequencies of EMS-induced stop codons and missense mutations are thought to be around 5 and 65%, respectively (McCallum et al., 2000).

In a first screen for leaf mutants (Fig. 2; Berná et al., 1999), we analyzed 46,159 M_2 seeds derived from 5,770 M_1 parental lines mutagenized with EMS at the seed stage by a commercial supplier (Lehle seeds; http://www.arabidopsis.com/). In the M_2 generation, we found 23.5% non-germinating seeds and 1.9% albino plants, as expected for an efficient mutagenesis. After visual inspection of individual M_3 plants, we selected 1,926 putative mutants (4.2% of the M_3 population) with leaves clearly different from those of Ler or that were abnormally arranged in the basal rosette. These mutants were initially assigned to three broad categories, corresponding to alterations in: (1) shape or size of the lamina, (2) margin configuration, or (3) structure of the basal rosette. Many putative M_2 mutants (853) yielded viable M_3 progeny but only 333 transmitted their mutant phenotype to the next generation in a Mendelian manner, with complete penetrance and only subtle variation in expressivity. Eighty-one mutants with altered rosette structure were discarded because they were not necessarily affected in the aspects of leaf development that were of interest to us. The inheritance patterns of the remaining 252 mutant lines were determined in backcrosses to Ler. The segregation ratios of the leaf mutant phenotypes were studied in their corresponding F_2 families, and the phenotypes of F_2 inbred descendants from the F_3 mutant plants were recorded. For subsequent genetic analyses, we selected 153 mutants whose phenotype remained unchanged in the M_2, M_3, and F_3 generations and was inherited as a monogenic trait.

A complementation test (Lewis, 1939) makes the assessment of allelism possible by crossing any two recessive mutants and analyzing the phenotype of their F_2 progeny. Despite the simplicity of this test, the number of crosses required to assign a large collection of mutants to a discrete number of complementation groups becomes enormous and time-consuming even for a moderate number of mutants. For instance, testing the allelism of 153 mutants requires 153/(2!(153-2)) = 11,628 crosses. To reduce this number, we defined 19 phenotypic classes (Fig. 3) by grouping mutants with similar characteristics and limited the crosses to mutants within the same class. This seemed reasonable because mutant alleles of the same gene are likely to yield similar phenotypes. The phenotypic classes and the corresponding genes were given Latin names descriptive of their most conspicuous leaf characteristics, inspired by Stearn’s Botanical Latin (1995). One exception was the Asymmetric leaves phenotypic class, whose name was already in use for mutants with leaf bilateral asymmetry and a slightly revolute leaf margin (Rédei and Hirono, 1964).

After completing 525 crosses, we assigned our 153 leaf mutants to 94 complementation groups. Since each M_1 parental was

<table>
<thead>
<tr>
<th>M_1 seeds</th>
<th>EMS</th>
<th>Fast neutrons</th>
<th>Time (weeks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M_1 plants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M_2 seeds</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M_2 plants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M_3 seeds</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M_3 plants x Ler plants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F_1 seeds</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F_1 plants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F_2 seeds</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F_2 plants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F_3 seeds</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F_3 plants</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2. Flow-chart for the isolation and genetic analysis of leaf mutants. The symbols and indicate growth and selfing, respectively. The time scale refers to each mutant. Tasks indicated in italics were outsourced to a company.
representations of the same parental group represented repeated isolations of the same mutation. Thus, after correcting for this fact, we conservatively estimated the number of different mutations to be 120. When classified according to their mode of inheritance, 113 of the mutations were recessive, six were semi-dominant, and only one was completely dominant. According to the number of alleles per gene, the 94 complementation groups were defined by one (76 complementation groups), two (11), three (six), or four (one) mutant alleles.

Assuming equal mutability of all genes, the number of mutant alleles per gene in a mutagenesis experiment is often thought to follow a Poisson probability distribution (Jürgens et al., 1991). To reach saturation, the number of mutants obtained should be high enough to render negligible the probability of a gene not being represented by any mutant allele. With an average of 1.27 (120/94) alleles per gene, the fraction of genes that mutate to viable leaf phenotypes but were missed by our screen can be estimated from the zero term of a Poisson distribution to be \( e^{-1.27} = 27.8\% \). Although relatively far from saturation, we identified a broad spectrum of leaf morphological alterations that appear in Arabidopsis thaliana after chemical mutagenesis, and they provided us with a precious collection of viable, fertile mutants that has helped us and other groups to dissect specific aspects of leaf development.

Physical mutagenesis with X-rays or fast neutrons normally creates deletions and chromosomal rearrangements (Shirley et al., 1992). In a second screen, we sowed 23,445 M\(_2\) seeds derived from 2,931 M\(_1\) parents irradiated with fast neutrons. We found 1% albino plants in the M\(_2\) population and 901 putative leaf mutants, 523 of which produced M\(_3\) seeds (Robles and Micol, unpublished results). In 232 of these mutants, the phenotype was transmitted to the M\(_3\) generation with complete penetrance and only subtle variation in expressivity. We selected 25 mutant lines with stable phenotypes and grouped them into five phenotypic classes. Complementation tests allowed us to conclude that they represented eight damaged genes (Robles and Micol, unpublished results).

The Arabidopsis Information Service (AIS) Form Mutants collection, which includes mutants isolated by G. Röbbelen and A. R. Kranz, was the largest morphological mutant collection available at that time. We selected 115 AIS form mutants for genetic analysis (Serrano-Cartagena et al., 1999). The leaf mutant phenotypes were inherited as recessive monogenic traits and were classified into 11 previously defined phenotypic classes (Berná et al., 1999), plus three new classes. These AIS mutants fell into 37 complementation groups. The most abundant class, Incurvata (icol), comprised 22 mutants with involute (curled upward) vegetative leaves, falling into ten different complementation groups. The study of double mutants involving alleles of five ICU genes allowed us to classify them into three groups based on the genetic interactions detected among them (Serrano-Cartagena et al., 2000).

In addition to experimentally induced mutants, we also used natural allelic variation to study the genetic basis of leaf morphology. We screened 188 natural accessions for variations in leaf architecture. These lines were grown under strictly controlled environmental conditions and assigned to 14 phenotypic categories based on petiole length, marginal serration, and overall lamina shape (Pérez-Pérez et al., 2002a). Accessions showing extreme and opposite variations were crossed and their progeny studied. Continuous variation was present in the F\(_2\) progeny, as expected for polygenic traits. Interestingly, leaf shape and size were among the characters initially studied by Mendel but were
finally discarded by him due to difficulties in assigning the observed variation to discrete classes (Lewis, 2001). Our results suggested that diverse leaf morphology in Arabidopsis thaliana results from natural allelic variation at quantitative trait loci (QTL). These observations prompted us to characterize leaf morphology in two recombinitant inbred line (RIL) populations, derived from Ler-0 x Col-4 and Ler x Cvi crosses, in an attempt to detect, map, and estimate the effects of the QTL responsible for the differences between natural accessions (Pérez-Pérez et al., 2002a; Juenger et al., 2005). We detected 16 and 13 QTL affecting highly correlated leaf morphological traits in juvenile and adult leaves, respectively, using the Ler x Col-4 RIL population, and three additional QTL for juvenile leaf traits using the Ler x Cvi RIL population. Finding QTL confirmed the multifactorial nature of the observed variation in leaf architecture among Arabidopsis thaliana natural accessions.

Leaf venation follows a complex net-like pattern in Arabidopsis thaliana. We also screened for natural variants in the venation pattern of the first vegetative leaves in 266 Arabidopsis thaliana accessions, finding two, Ba-1 and Ei-5, with abnormal vascular patterns. Ei-5 has an extremely simple pattern that we named Hemivenata and is inherited as a monogenic recessive trait (Candela et al., 1999). Finally, to test if vascular patterning is regulated independently of whole leaf morphogenesis, we screened for mutants with abnormal vein patterns and normal leaf shape and size. More than 6,000 EMS-mutagenized M<sub>2</sub> seeds and 17,313 T<sub>4</sub> plants derived from 1,800 accessions, finding two, Ba-1 and Ei-5, with abnormal vascular patterns. Ei-5 has an extremely simple pattern that we named Hemivenata and is inherited as a monogenic recessive trait (Candela et al., 1999).

Allelism tests were subsequently performed allowing us to assign map positions to 97 of these genes (Robles and Micol, 2001). Allelism tests were subsequently performed only for closely linked mutations.

A major step in the genetic characterization of a mutant is the molecular identification of the corresponding gene affected by the mutation. Mutations tagged by insertional mutagenesis often facilitate the recovery of the genomic sequences flanking the insertion element. A drawback of this approach is that insertion mutagenesis usually yields null alleles that might be lethal if homozygous. In contrast, chemical mutagenesis is more efficient and usually produces viable hypomorphic and antimorphic alleles (Noveroske et al., 2000).

Map-based cloning is a classic approach for the molecular identification of a gene when a molecular tag is not available. It involves genotyping the individuals of an F<sub>2</sub> mapping population to identify recombination events that occurred in F<sub>1</sub> hybrids derived from an outcross involving the mutant of interest (Fig. 4). With the completion of the genome sequence and the availability of enough molecular markers to cover the entire genome, map-based cloning emerged as a powerful approach to identifying genes altered by mutation in many experimental organisms, including plant species other than Arabidopsis thaliana (Bortiri et al., 2006). The number of polymorphic markers and F<sub>2</sub> individuals needed to detect unambiguous linkage between the mutation under study and a linked genetic marker is proportional to genome size. In Arabidopsis thaliana, with a genome of 500 cM (125 Mb) distributed on five chromosomes, a minimum of 17 markers regularly spaced every 30 cM and a mapping population of 15-20

Map-based cloning is a classic approach for the molecular identification of a gene when a molecular tag is not available. It involves genotyping the individuals of an F<sub>2</sub> mapping population to identify recombination events that occurred in F<sub>1</sub> hybrids derived from an outcross involving the mutant of interest (Fig. 4). With the completion of the genome sequence and the availability of enough molecular markers to cover the entire genome, map-based cloning emerged as a powerful approach to identifying genes altered by mutation in many experimental organisms, including plant species other than Arabidopsis thaliana (Bortiri et al., 2006). The number of polymorphic markers and F<sub>2</sub> individuals needed to detect unambiguous linkage between the mutation under study and a linked genetic marker is proportional to genome size. In Arabidopsis thaliana, with a genome of 500 cM (125 Mb) distributed on five chromosomes, a minimum of 17 markers regularly spaced every 30 cM and a mapping population of 15-20

High-throughput gene mapping as an alternative to massive complementation analysis

From our screens for leaf mutants using either EMS or fast neutrons as mutagens, and those mutants obtained by other authors, we ended up with 94, 8, and 37 genes, respectively (see above). As an alternative to massive crossing for testing the allelism of mutations affecting leaf development, we developed a high-throughput mapping procedure based on simultaneous PCR coamplification, fluorescent semiautomated detection, and sizing of several microsatellite markers (Ponce et al., 1999), which

Preparation of a mapping population

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Polymorphic ecotype</th>
<th>F&lt;sub&gt;1&lt;/sub&gt; (wild type)</th>
<th>F&lt;sub&gt;2&lt;/sub&gt;</th>
</tr>
</thead>
</table>

Fig. 4. Schematic representation of our high-throughput gene mapping method. To assess the linkage of a given phenotype to polymorphic molecular markers, the mutant line was first outcrossed to the Columbia-0 accession and the F<sub>2</sub> plants displaying the recessive phenotype were harvested for their DNA extraction. From each individual F<sub>2</sub> plant, DNA was amplified in a single-tube multiplex PCR as described in Ponce et al. (1999, 2006). The amplification products were electrophoresed in an ABI PRISM 3100 and sized with fragment analysis software.
F₂ plants are required to significantly detect linkage between a mutation under study and at least one marker (Ponce et al., 1999). Several high-throughput but low-resolution mapping approaches have been developed in Arabidopsis thaliana. PCR-based mapping with simple sequence length polymorphism (SSLP) markers, either alone (Ponce et al., 1999) or combined with small insertions and deletions (Jander et al., 2002; Ponce et al., 2006), with amplified fragment length polymorphism (AFLP) markers (Peters et al., 2004), or with high-density oligonucleotide array hybridizations (Spiegelman et al., 2000; Gong et al., 2004). The last approach has been useful for mapping moderately large deletions such as those induced by physical mutagens.

Based on our experience, successful positional cloning of Arabidopsis thaliana genes starts with low resolution mapping of a mutation using about 50 F₂ mutant homozygotes (Ponce et al., 1999; Robles and Micol, 2001). Subsequently, if no obvious candidate gene is found in the genomic region, the mutation is fine mapped using 400-1000 additional F₂ individuals. Once the F₂ mapping population is ready, less than 2 weeks are required for cloning, including 100-200-kb candidate interval (Ponce et al., 2006).

**Genes cloned, molecular nature of their products, and processes in which they are involved**

The map positions obtained with the high-throughput method outlined above (Ponce et al., 1999; 2006) have led to the positional cloning of genes identified in our screens (Berná et al., 1999; Serrano-Cartagena et al., 1999; Robles and Micol, unpublished). This is the case for 25 of the 97 mutations we initially mapped, including 86 induced by EMS, nine taken from the AIS published, in some cases in collaboration with other groups, the cloning and functional characterization of 13 genes that affect leaf development (Table 2). The first two genes that we cloned, UCU1 and UCU2, encode proteins involved in the signal transduction of brassinosteroids (BR), steroid hormones that control growth processes such as cell expansion, vascular differentiation, and reproductive development. The ucu1 and ucu2 mutants resemble BR-biosynthesis or BR-perception mutants (Fig. 3S): dwarf plants with roots, hypocotyls, stems, and fruits shorter than in the wild type (Pérez-Pérez et al., 2002b; 2004). UCU1 [also called BRASSINOSTEROID INSSENSITIV2 (BIN2); Li and Nam, 2002] is a member of the SHAGGY/GSK3 family of intracellular protein kinases, which negatively regulate several signal transduction pathways in metazoans, such as those triggered by insulin or the Wingless/Wnt signal (Cohen and Frame, 2001). In Arabidopsis thaliana, UCU1 may act as a negative modulator of the BR transduction pathway and likely in the crosstalk of the BR and auxin signaling pathways (Pérez-Pérez et al., 2002b). UCU2 encodes a peptidyl-prolyl cis/trans-isomerase of the FK506-binding protein family involved in protein activation. In addition to BR signaling, UCU2 might also be required for the transport of auxin and, like UCU1, UCU2 plays an important role in the overlap between BR and auxin signaling pathways during leaf expansion (Pérez-Pérez et al., 2004).

In collaboration with Mieke van Lijsebettens’ group, RON2 was identified as the same gene as LEUNIG (LUG; Conner and Liu, 2000), which encodes a WD40 repeat protein similar to transcriptional co-repressors in yeast, Drosophila, and humans (Cnops et al., 2004). RON2 seems to act at the late stages of leaf development and may control leaf growth through restriction of cell expansion. Perturbation of RON2 activity causes an increase in leaf cell size, as observed in the ron2 mutants, which display wider laminae and larger petioles than the wild type (Cnops et al., 2004). ELO1, ELO2, and ELO3 are orthologs of the yeast genes ELP4, ELP1, and ELP3, respectively, whose products form part of the Elongator complex (Nelissen et al., 2005), a histone acetyl transferase that co-purifies with RNA polymerase II and facilitates the passage of the polymerase along the template during transcription elongation (Otero et al., 1999; Hawkes et al., 2002). The ELO4 gene is the same gene as DEFORMED ROOTS AND LEAVES1 (DRL1), an ortholog of the yeast TOT4/KTI12 gene.

**Table 2**

<table>
<thead>
<tr>
<th>Mutants</th>
<th>Genes affected</th>
<th>Alternative gene names</th>
<th>Function of the gene cloned</th>
<th>Developmental process</th>
</tr>
</thead>
<tbody>
<tr>
<td>ulcerata (uc1)</td>
<td>UCU1</td>
<td>BIN²</td>
<td>Signal transduction of brassinosteroids and auxin</td>
<td>Polar cell expansion</td>
</tr>
<tr>
<td>ulcerata2 (uc2)</td>
<td>UCU²</td>
<td>TWD², TWD1⁴</td>
<td>Signal transduction of brassinosteroids and auxin</td>
<td>Polar cell expansion</td>
</tr>
<tr>
<td>rotundifolia (ron2)</td>
<td>RON²</td>
<td>LUG³</td>
<td>Transcriptional co-repressor</td>
<td>Restriction of leaf cell expansion</td>
</tr>
<tr>
<td>elongata (elo1-elo4)</td>
<td>ELO1-ELO3⁷</td>
<td>ELO4⁸</td>
<td>Elongation of nuclear gene transcription</td>
<td>Leaf lateral expansion</td>
</tr>
<tr>
<td>reticulata (rt)</td>
<td>RE⁰</td>
<td>LCO¹, VEN²¹</td>
<td>Unknown function</td>
<td>Differential contribution of the epidermal and internal tissues to leaf organogenesis</td>
</tr>
<tr>
<td>incurvata (icu4)</td>
<td>ICU⁴⁴</td>
<td>CNA¹</td>
<td>Transcription factor</td>
<td>Meristem activity</td>
</tr>
<tr>
<td>scabra3 (scab3)</td>
<td>SCA³⁴</td>
<td>Choroplast gene transcription</td>
<td>Relationship between plastid biogenesis and leaf organogenesis</td>
<td></td>
</tr>
<tr>
<td>hemivena (hve)</td>
<td>HVE¹⁵</td>
<td>CAND1⁴⁵, ETA²¹</td>
<td>Ubiquitination</td>
<td>Venation pattern formation</td>
</tr>
<tr>
<td>angustata (ang4)</td>
<td>ANG⁴⁹</td>
<td>HUB¹⁹</td>
<td>Histone ubiquitination</td>
<td>Cell cycle regulation</td>
</tr>
<tr>
<td>incurvata2 (icu2)</td>
<td>ICU²¹⁵</td>
<td>DNA replication</td>
<td>Chromatin-mediated cellular memory</td>
<td></td>
</tr>
</tbody>
</table>
putatively involved in the regulation of the Elongator complex (Nelissen et al., 2003). Perturbation of the ELO genes causes a reduction in overall cell division rate that explains the delay in germination and the retarded seedling and primary root growth observed in elo mutants (Fig. 3H). The product of HISTONE MONOUBIQUITINATION1 (HUB1; initially named ANGUSTA4 [ANG4]), is a functional homolog of yeast and human histone H2B-monoubiquitinating BRE1 E3 ligases (Fleury et al., 2007; Liu et al., 2007), which are associated with chromatin modifications and transcriptional activation. Mutations in HUB1 affect plant growth and cause reduced plant size, narrow leaves (Fig. 3I), and a reduction in root growth rate, as a consequence of a block in the G2-to-M transition and a prolonged duration of the cell cycle, as also revealed by microarray analyses (Fleury et al., 2007). HUB1 controls gene activation and cell cycle regulation in leaf cells through chromatin modifications.

Together with Antonio Martínez-Laborda’s group, we found that the ICU4 gene, also known as CORONA (CNA; Green et al., 2005), encodes the HD-ZIPIII transcription factor ATH15 (Ochando et al., 2006). The transcripts of ICU4 and other HD-ZIPIII genes have a microRNA (miRNA) complementary site and are under post-transcriptional regulation. The semidominant icu4-1 and icu4-2 alleles carry identical nucleotide substitutions at the miRNA165/166 complementary site of ICU4. These mutations lead to defective cleavage and accumulation of ICU4 transcripts in semidominant icu4 mutants. Phenotypic studies revealed that ICU4 plays an adaxializing role in leaves, and is required for shoot meristem function and for the patterning of shoot vascular bundles (Ochando et al., 2006).

The scabra3 (sca3) mutant, with characteristic pale green leaves, shares with re mutants a marked decrease in mesophyll cell number. The SCA3 protein corresponds to the nuclear encoded plastid-targeted RNA polymerase RpoTp, which controls the expression of a subset of plastid genes (Hricová et al., 2006). Strong loss-of-function sca3 alleles affect the number, size and morphology of chloroplasts, and perturb the expression of both plastid and nuclear genes, many encoding chloroplast-targeted proteins (Hricová et al., 2006). Our analyses of sca3 mutants demonstrate the intertwining of plastid biogenesis and leaf organogenesis.

The hemivenenta-1 (hev-1) recessive allele was isolated in a search for natural variants of the Arabidopsis thaliana leaf venation pattern (Candela et al., 1999). hve-1 reduces the complexity of the venation pattern in leaves and cotyledons. The HVE1 gene encodes the Arabidopsis thaliana ortholog of the mammalian CAND1 protein (Alonso-Peral et al., 2006), which binds to un neddyled CULLIN1 (CUL1), a protein that forms part of the SCF ubiquitin ligase complexes that regulate target protein degradation through the 26S proteasome. Similar venation patterns in auxin resistant1 (axr1), a mutant insensitive to auxin, and hve mutants indicate that auxin signaling mediated by ubiquitination is required for venation patterning in Arabidopsis thaliana leaves (Alonso-Peral et al., 2006).

INCURVATA2 (ICU2) encodes the putative catalytic subunit of the DNA polymerase α of Arabidopsis thaliana (Barrero et al., 2007). Genetic and molecular interactions between chromatin remodeling factors and components of the DNA replication machinery have been identified in yeast. The weak icu2-1 allele causes early flowering, leaf incurvature, and homeotic transformations in flowers, traits that can be attributed to derepression of flowering time, floral meristem, and floral organ identity genes, respectively. Our genetic analyses indicate that ICU2 interacts with TERMINAL FLOWER2, the ortholog of HETEROCHROMATIN PROTEIN1 of animals and yeasts (Takada and Goto, 2003), and with the Polycomb group (PcG) gene CURLY LEAF (Goodrich et al., 1997). Our results suggest the existence in plants of a cellular memory mechanism that involves DNA polymerase α, as proposed by previous authors for the fission yeast.

Contrary to all genes discussed above, no clear function could be assigned to the RETICULATA (RE) gene (González-Bayón et al., 2006), named LOWER CELL DENSITY1 (LCD1) by Barth and Conklin (2003). The RE gene is ubiquitously expressed, and the RE protein has been identified in the chloroplast proteome (Dunkley et al., 2006). The re mutant was initially isolated by Rédei and Hirono (1964), and used as a classic phenotypic marker for mapping for many years. Our histological characterization of re mutants shows a significant reduction in mesophyll cell density but little effects on leaf shape, indicating that RE is normally required for the development of internal leaf tissues (González-Bayón et al., 2006). These results were obtained in collaboration with the group of Kevin Pyke and suggest that the epidermis plays a major role in controlling leaf shape in Arabidopsis thaliana, whereas the division and expansion of the internal tissues appear to be secondary in this process, as in tobacco leaves (Avery, 1933). Furthermore, re and wild-type leaves have similar shapes, suggesting that the epidermis plays a major role in controlling leaf shape.

Concluding remarks

From our EMS screen in the crucifer Arabidopsis thaliana, we estimate that about 130 genes [94×100/(100-27.8)] yield morphological leaf phenotypes and hence are required for normal leaf ontogeny. However, this value underestimates the real number of genes because it is based on an average number of alleles per gene (1.27) that did not consider genes not identified in the screen. Estimates that take them into account are expected to be more accurate. By applying the formulae of Rice (1973), Meneely and Herman (1979), and Engels (1981) to our data, we obtain values of 357, 279, and 233±39 genes, respectively. An immediate conclusion from these results is that a significant fraction of the genome must work cooperatively to produce a leaf. These values only refer to those genes that can mutate to viable leaf shape phenotypes, and represent a lower limit that does not take into account mutations with pleiotropic effects that prevent their transmission, such as those that cause sterility or are lethal. For example, genes essential for developmental decisions in the embryo may also act at later stages, but the study of their postembryonic functions will only be possible using special techniques, such as clonal analysis.

Mutations in genes required for the establishment of shoot meristem and leaf identities have been extensively studied in plants. In maize, ectopic expression of class I KNOX genes in dominant Knot ted1 and Rough sheath1 mutants causes abnormal cell proliferation and differentiation patterns (Hake et al., 1989; Schneeberger et al., 1995). The Rough sheath2 gene of maize (Tsiantis et al., 1999) and its orthologs PHANTASTICA (PHAN; Waite et al., 1998) and ASYMMETRIC LEAVES1 (AS1;
participate in a given biological process is necessary to understand how the process works. A genetic dissection normally begins with the isolation of mutations perturbing the normal course of development, to identify the genes at work in the process under study. A high-throughput genetic screen for mutants with altered leaf size and shape aiming to understand the mechanisms of leaf-size control was recently published (Horiguchi et al., 2006a; 2006b). A comparative study between these two large collections of mutants should help to identify most of the genes involved in leaf growth and development in Arabidopsis thaliana. The positional cloning of untagged mutations in model organisms such as Arabidopsis thaliana has proved a very fruitful approach even though it occasionally yields genes with very limited novelty or involved in totally unexpected functions. We have shown that the range of leaf phenotypes identified in our laboratory results from perturbation of a variety of unrelated gene functions. Combined with elegant work by other authors, our results have begun to reveal the outstanding complexity of leaf development. The molecular identification of additional genes uncovered by mutation will provide a better picture of how leaves develop and is expected to help explain the immense diversity in leaf shape and leaf size present in nature.

Acknowledgments
The authors wish thank S. Gerber, J.M. Serrano, V. García, and T. Trujillo for technical assistance during all these years, and M. van Lijsebettens, V. Aguilera, A. Mollá-Morales, A. Ferrández-Ayela, R. Sarmiento-Mafús, D. Hernández-Romero, and S. Rubio-Díaz for comments on the manuscript. This work was supported by grants from the Ministerio de Educación y Ciencia de Spain [PB81-0749, ACP95-0191, PB95-0685, PB98-1309, BMC2002-02840, BMC2005-01301, BIO2008-04075 and CSD2007-00057 (TRANSPLANTA)] and the European Commission [HPRN-CT-2002-02067 (DAGOLIGN) and LSHG-CT-2006-037704 (AGRON-OMICS)].

References
CHENG, Y., DAI, X. and ZHAO, Y. (2004). AtCAND1, a HEAT-repeat protein that...


Leaf mutant screens in Arabidopsis 1633


