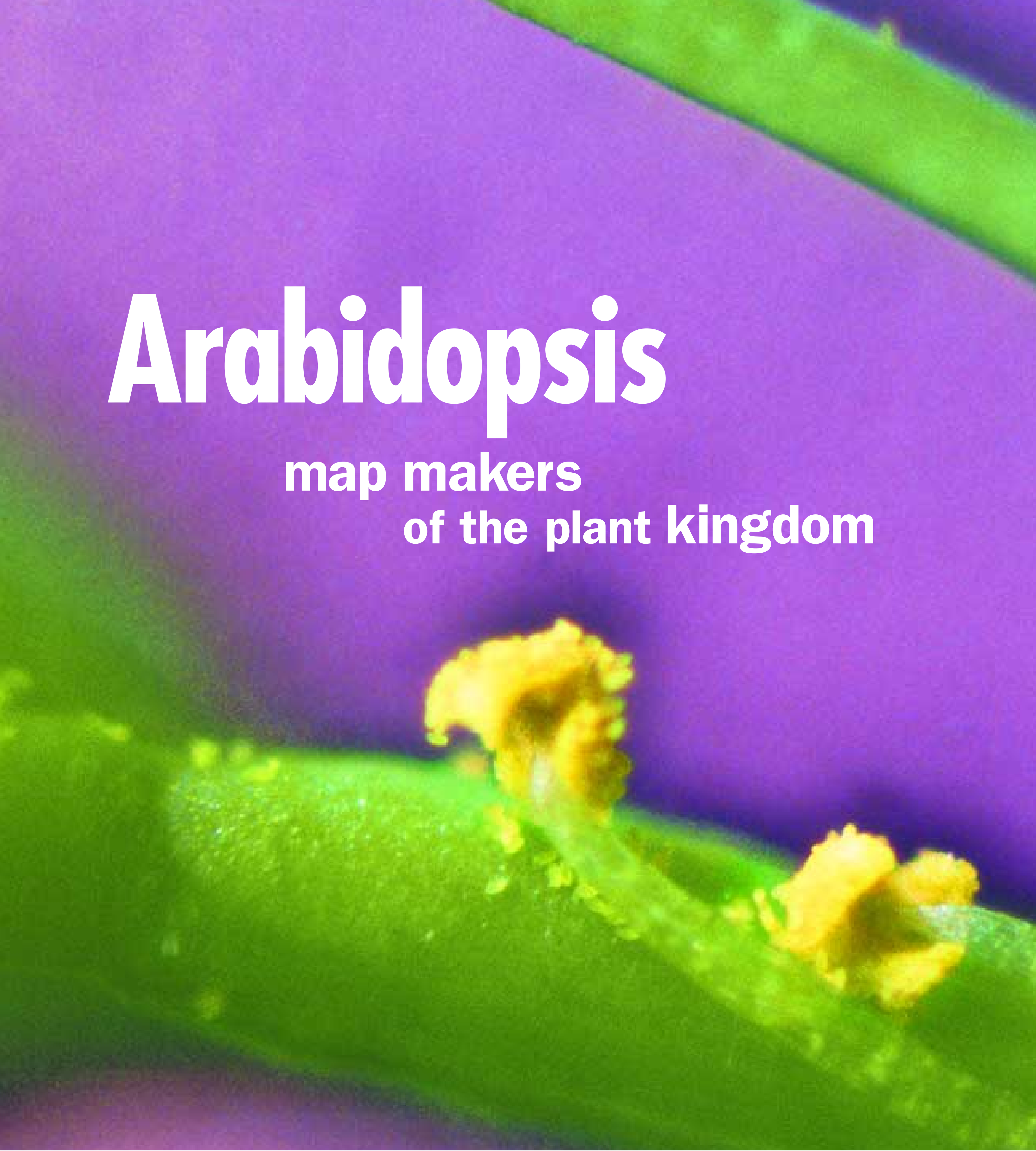
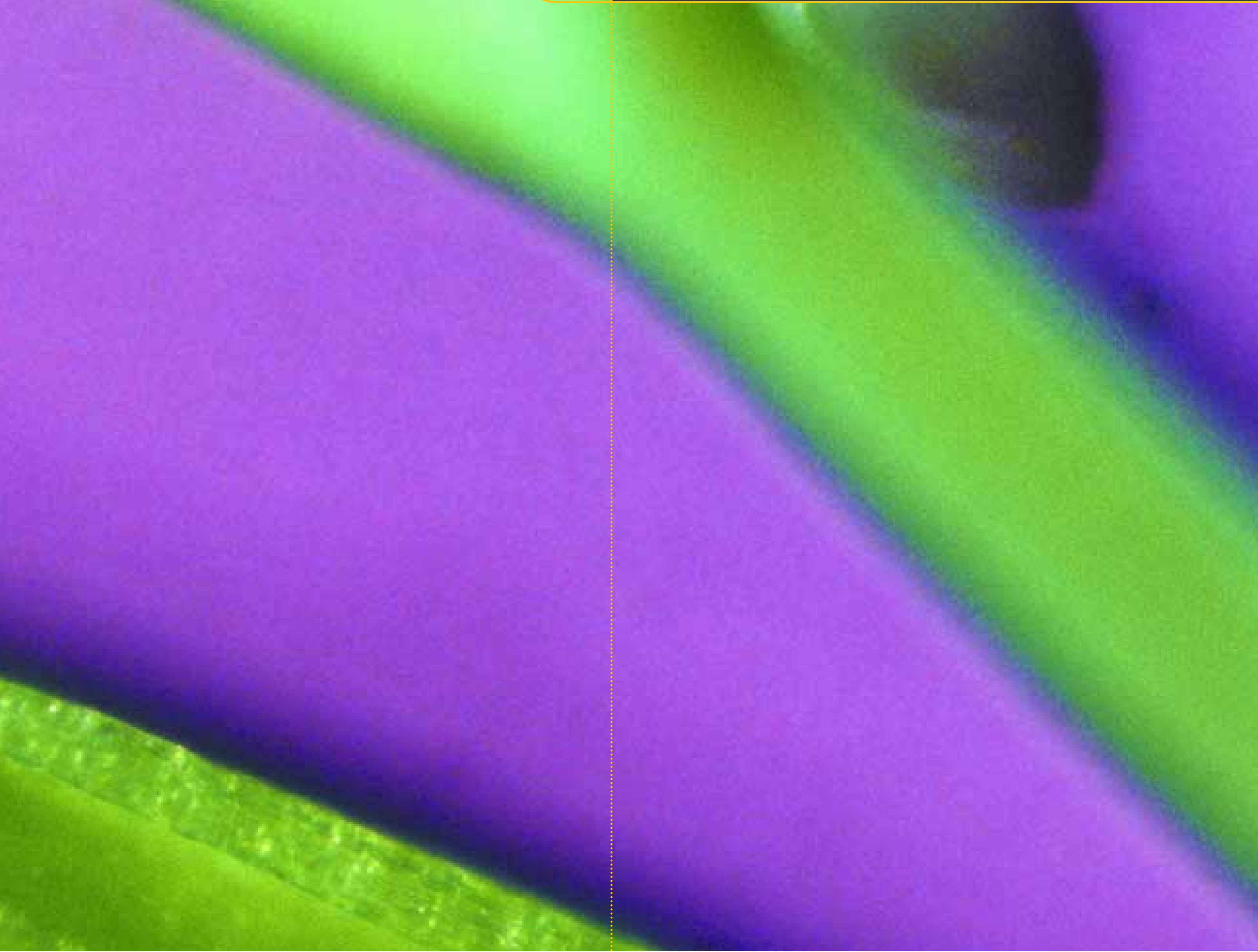
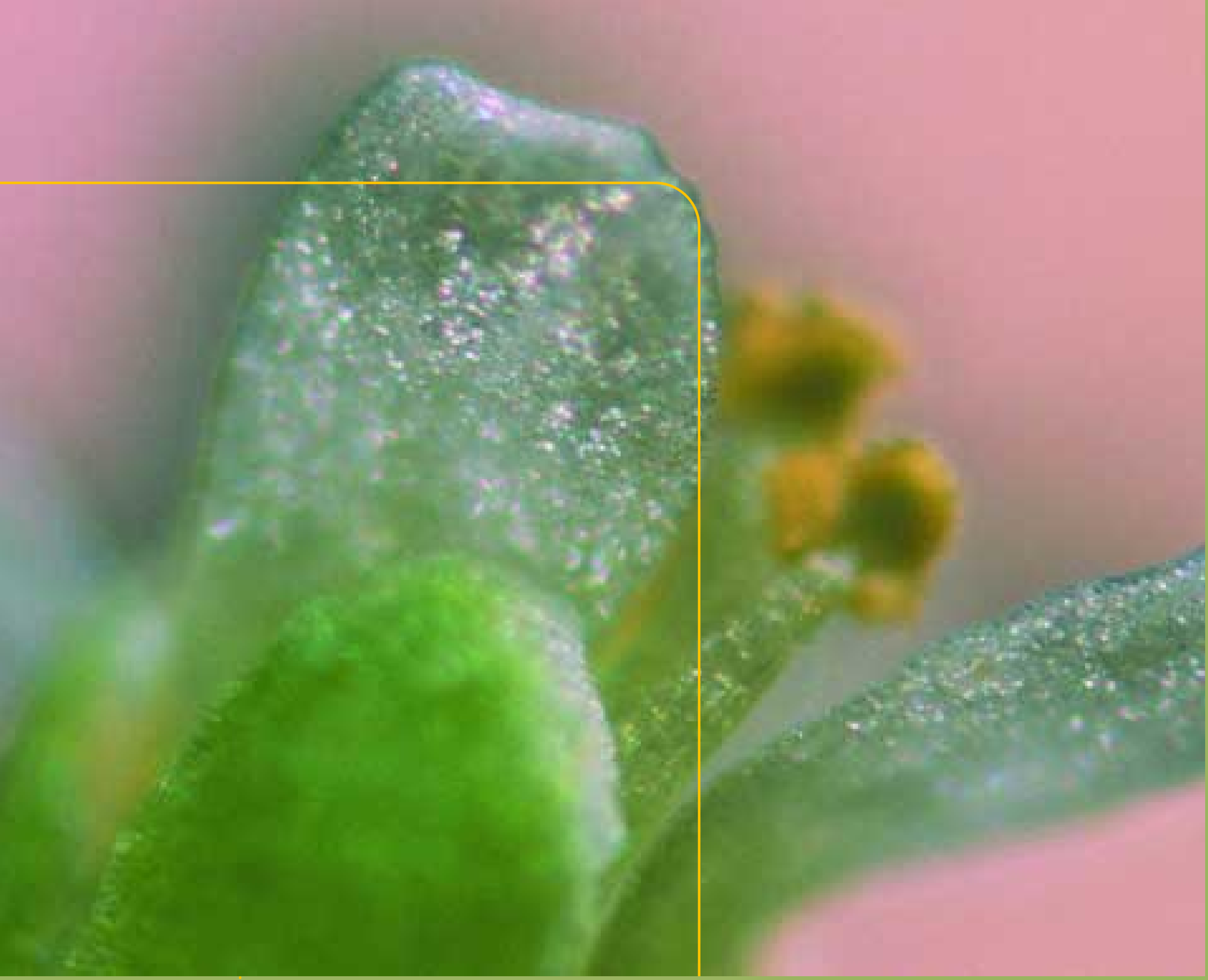
Arabidopsis

map makers of the plant kingdom



With NSF support, biologists today are mapping all of the genes of a model organism—identifying the location and function of each gene. They have already made fundamental discoveries that may lead to the development of more beneficial crops and forest products.





Arabidopsis thaliana is a small, flowering mustard plant that has become the subject of intense study by scientists around the world. It has many characteristics of an ideal experimental system—a model organism for elucidating the biology of flowering plants. Recognizing the promise of *Arabidopsis,* NSF began working with leaders in plant biology in the 1980s to cultivate a spirit of cooperation and to encourage the use of the model plant in research. In 1990, NSF launched a multi-agency, multinational project to identify all of the genes in *Arabidopsis* by location and function—in other words, to create a genetic road map to flowering plants. The collegial *Arabidopsis* research community now expects to complete the sequencing of the plant's genome by the end of 2000—several years ahead of schedule. By August 1999, nearly 70 percent of the genome sequences for *Arabidopsis* had been deposited in the public database, GenBank. Six months later, scientists reported the complete DNA sequences of two of the five chromosomes of *Arabidopsis*. Major discoveries have been made about the mechanisms by which genes regulate flower development, disease resistance, response to adverse conditions in the environment, and numerous other aspects of plant biochemistry and physiology. Commercial applications under development include trees with accelerated blooming, biodegradable plastics grown in crops, and genetically engineered vegetable oil with reduced polyunsaturated fat.

A Rose Is a Rose Is a Mustard Weed

There are approximately 250,000 different species of flowering plants, all believed to derive from a common ancestor. While plants have adapted to a multitude of terrains, climate conditions, and selective breeding efforts over the millennia, the process of evolution ensures that they remain related in fundamental ways. At the molecular level, for example, what causes a rose bush to flower is not terribly different from what occurs in a radish plant. Other characteristics also appear to be similar across species, such as the fruit-ripening process and the internal clock that tells plants when to open their pores in anticipation of daylight. In fact, the physiology and biochemistry of plants display such uniformity across species that one can say, without too much exaggeration: When you've seen one flowering plant (at the molecular level), you've seen them all.

This essential truth has altered the course of study in plant biology, a field once dominated by research into individual crops, such as corn or wheat. Today, plant biology has its own model organism, the flowering mustard plant *Arabidopsis thaliana*; consequently, research in the field now resembles other types of broad basic research, such as that done on bacteria or animals.

Considered a weed because it is uncultivated and grows in profusion, *Arabidopsis* nonetheless engages the attention of a global research community. The researchers, and agencies such as NSF that support them, expect that by analyzing the structure and functions of *Arabidopsis*, they are laying the groundwork for analyzing most other plant species.

The project has greatly accelerated the practical application of basic discoveries in agriculture and forestry. Genetically engineered species are beginning to appear, and many believe they signal the beginning of a revolution in plant breeding.

In one such area of discovery, scientists have identified genes involved with the regulation and structural forms of flowers. Knowledge of these genes has made possible the genetic engineering of plants other than *Arabidopsis*. For example, aspens normally flower only after they have attained a height of 30 feet, which can take up to twenty years. A genetically transformed aspen, however, flowered in only six months, when it was just 2 inches tall. Commercial tree growers have

Researchers are studying the inherited characteristics of *Arabidopsis*. The NSF-funded genome research project to map *Arabidopsis* will yield important information about how flowering plants interact with their environments.

This scanning electron micrograph shows an *Arabidopsis* plant in bloom, highlighting the emergence of the plant's flower. The image provides details about the flowering process that help researchers follow the earliest events in the development of an *Arabidopsis* flower.

always wanted to control the timing of floral and fruit production, as well as the closely related reproductive cycle. The technology is also being tested in fruit and timber trees.

In another research initiative, health concerns over saturated fat and hydrogenated vegetable oils are motivating a search for edible oils that pose no threat to human health. The pathways by which plants produce edible unsaturated oils have been elucidated and the responsible desaturase genes cloned from *Arabidopsis*. The *Arabidopsis* genes were used to identify the corresponding genes in soybeans and other crop plants, whose oils account for approximately one-third of the calories in the American diet. At present, most plant oils are chemically hydrogenated to keep them from turning rancid. The availability of the desaturase genes raises the possibility that nutritionally desirable edible oils can be produced from plants without the need for chemical modification. Agrichemical producers have begun field trials with modified soybeans and other plant species.

Inside the Little Green Factories

Plant breeding became a science around the turn of the twentieth century, thanks to Austrian scientist and mathematician Gregor Mendel. His studies of heredity in peas enabled him to draw conclusions about gene functioning by observing how the characteristics of parents showed up in generations of offspring. While adopting increasingly sophisticated techniques, plant breeders continued to improve crops in traditional ways, crossing the current stock with germplasm containing useful new characteristics. The success of the outcomes depended on the skill and judgment of the breeder in selecting plants to cross.

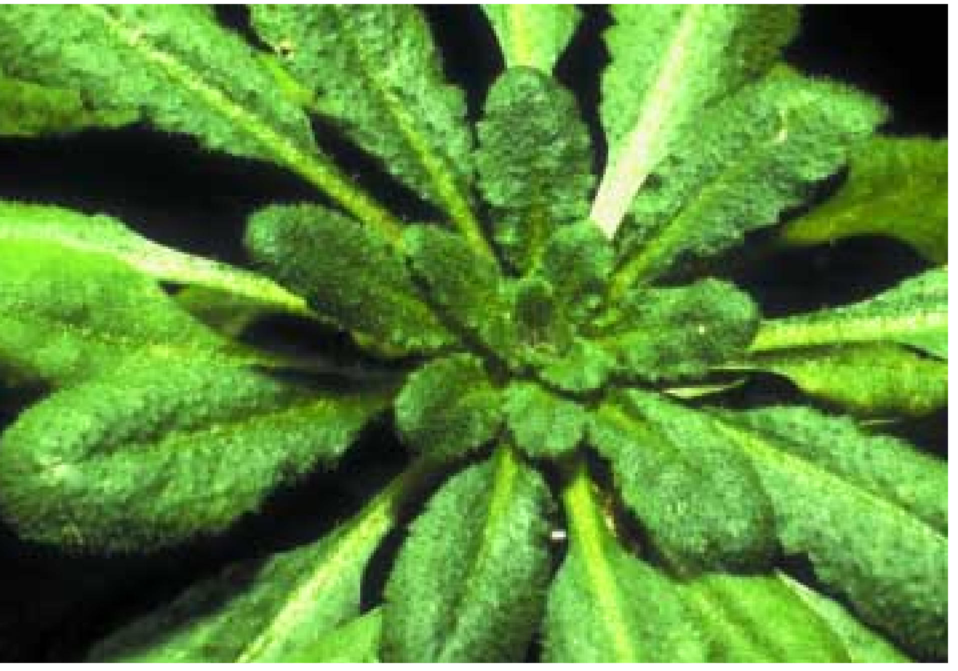
Scientists, meanwhile, sought to understand the underlying genetic mechanisms that induced plants to express inherited characteristics in certain ways. With advances in plant tissue culture techniques, biologists were able to produce novel hybrids and study them under controlled laboratory conditions. Of particular interest were plant characteristics that might potentially be modified in ways advantageous to humans. One scientist described plants as the "little green factories" that produce food, fibers, housing materials, and many pharmaceuticals, as well as the oxygen necessary for terrestrial life.

How to Make a Flower

Elliot M. Meyerowitz of the California Institute of Technology in Pasadena was one of the first molecular biologists to receive an NSF grant to study *Arabidopsis* genetics. His work on the development of flowers illustrates how the methods of scientific inquiry employed in molecular biology can unlock the secrets of plant life.

Flowers are made up of four concentric whorls. Surrounded by tough, protective structures called sepals, the petals themselves surround the male and female sex organs, respectively called stamens and carpels. Three types of genes control how the whorls develop, and by looking at flowers that lacked some genes, Meyerowitz's lab discovered that if only type A genes are active, a cell knows to become part of a sepal. With A and B genes switched on, the cell turns into part of a petal. Together, genes B and C direct a cell into a stamen, and C alone, into a carpel.

Meyerowitz's work has broad applicability. Fully 80 percent of the world's food supply is made up of flowers or flower parts: fruit, grains, or seeds. While genetically engineered flowers may have limited commercial value, the same formulas may one day be used to tailor food crops to the requirements of humankind.



This *Arabidopsis* plant, grown under short-day conditions (eight hours of light/sixteen hours of dark), shows how many leaves are produced during the vegetative phase of the plant's approximately five-week life cycle. After the transition from the vegetative to reproductive phase, the plant produces flowers.

Knowledge about genetics grew rapidly during the 1960s and 1970s, and certain characteristics became recognized as central to all organisms: bacteria, animals, plants, and humans. For example:

- Organs develop and function as they do because of the way different combinations of genes express themselves in the form of proteins produced within cells. The instructions that tell proteins to form a blood cell, a brain cell, or a flower petal are all contained within the genome, and insofar as is now known, in the chemical composition of the deoxyribonucleic acid, or DNA, in a particular gene sequence along a chromosome.
- When a gene or its constituent nucleotides undergoes a sudden random change, known as mutation, the result is an abnormality in the affected cells. Mutations that render an organism better able to cope with its environment are the raw material that natural selection acts on. Many of the successful mutations of an organism's ancestors, and possibly a mutation or two of its own, are reflected in the organism's genetic composition, or genome.
- The genome responds to environmental forces, such as the supply of essential nutrients, to produce an organism's observable characteristics, or phenotype.
- Through recombinant DNA technology, or genetic engineering, it is possible to create new strains of organisms with DNA containing the exact genes desired from different sources.

Barbara McClintock's work identifying mobile genes in corn, for which she received a Nobel Prize in 1983, provided molecular biologists with the tools necessary for the development of plant transformation. Despite the essential role that plants play in human existence, much less time and energy had gone into studying the genetic functioning of plants than of bacteria or animals —or humans. A major obstacle was the large and unwieldy mass of genetic material found in most crop plants, which were the primary subjects of scientific research.

This obstacle was a big one. A scientist who wants to find the genetic source of a mutation, such as resistance to a particular disease, has to examine the cells where the mutation was expressed and connect the genetic information there back to the DNA. The technologies for identifying and isolating genes, sequencing them, cloning them (making large numbers of exact reproductions), and determining their functions are complex, labor-intensive, and expensive. To apply these techniques to plants, molecular biologists needed a plant whose genome was of manageable size.

Golden Age of Discovery

"We are now in a 'golden age' of discovery in plant biology. Problems that have been intractable for decades are yielding to the application of modern methods in molecular and cellular biology. The formula for much of this success is conceptually simple: Isolate a mutation that affects the process or structure of interest, clone the gene, find out where and when it is expressed, where the gene product is located, what it does, and what it interacts with, directly or indirectly . . . . Although it is not necessarily easy, any gene that can be marked by a mutation can be cloned. This is a qualitatively different situation from anything that has ever before existed in plant biology."

—From *Arabidopsis*, E. Meyerowitz and C. Somerville, eds., Cold Spring Harbor Press, 1994.

Increasingly, they converged on *Arabidopsis thaliana,* a weed of the mustard family that has one of the smallest genomes of any flowering plant. It is estimated that 20,000 to 25,000 genes are arrayed on only five chromosomes, with little of the puzzling, interminably repetitious DNA that frustrates efforts to study most plants. *Arabidopsis* is compact, seldom exceeding about a foot in height, and it flourishes under fluorescent lights. All of these characteristics enable scientists to raise it inexpensively in laboratories. During its short life cycle, this mustard weed produces seeds and mutants prodigiously. It can be transformed through the insertion of foreign genes and regenerated from protoplasts, plant cells stripped of their cell walls. For all of its superior properties, *Arabidopsis* is typical of flowering plants in its morphology, anatomy, growth, development, and environmental responses, a kind of "everyman" of the plant world. In short, *Arabidopsis thaliana* is a biologist's dream: a model plant.

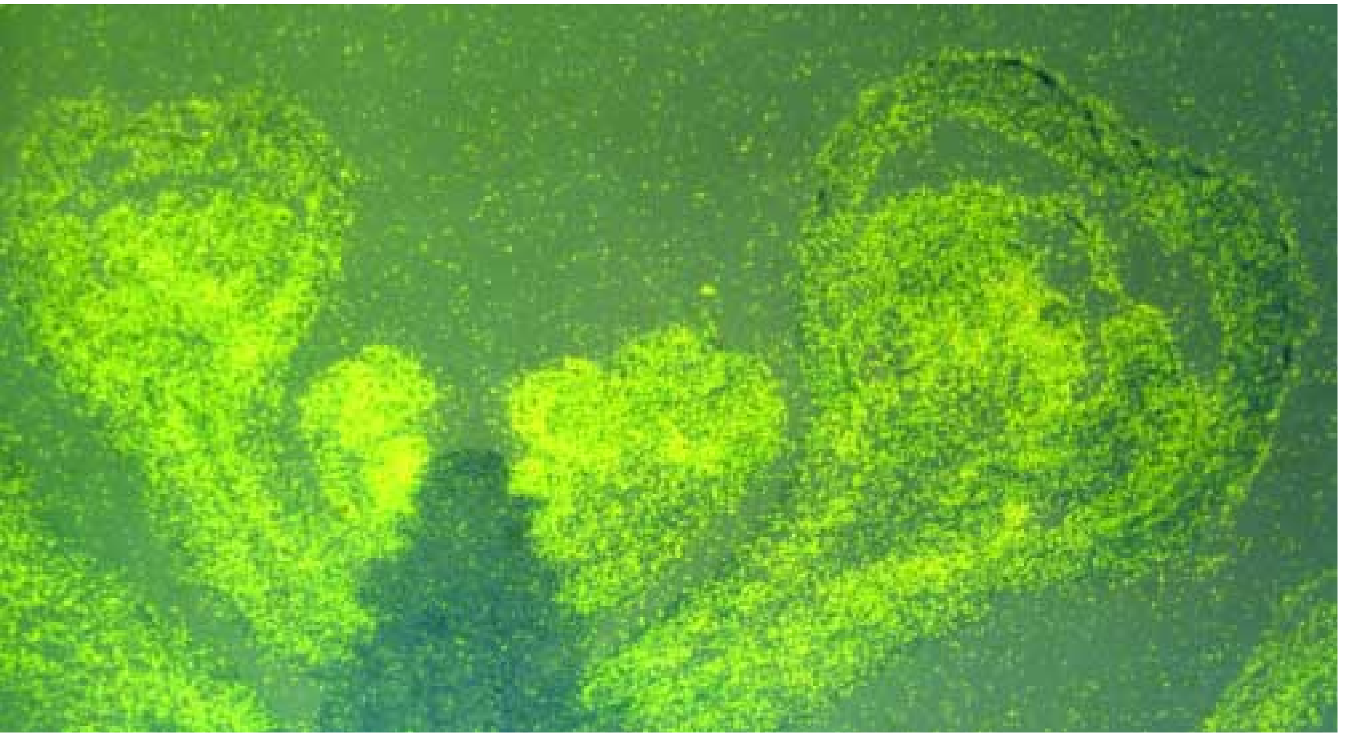
Researchers around the world are using the unassuming mustard weed to unlock the secrets of the plant world—secrets with many potential benefits.

NSF Helps Launch the New Biology

Arabidopsis began to intrigue not only plant biologists, but also scientists who formerly specialized in bacteria or fruit flies. As laboratories around the world undertook *Arabidopsis* projects, the stock of available mutants grew and new techniques were developed for gene cloning. Scientists began making breakthrough discoveries. And NSF undertook to advance *Arabidopsis* research even more rapidly—first through a series of workshops and then by launching a long-range plan in 1990 for the Multinational Coordinated *Arabidopsis thaliana* Genome Research Project. The project's steering committee, made up of scientists from eight countries, announced a collaborative agreement within the international community to pursue the goal of understanding the physiology, biochemistry, and growth and developmental processes of the flowering plant at the molecular level.

"I see the NSF program people as scientific collaborators," said Chris Somerville, director of the Plant Biology Department at Carnegie Institution of Washington in Stanford, California, and with Elliot Meyerowitz, co-author of the leading research compendium on *Arabidopsis*. "[NSF] sensed something happening in the community that the individual scientist didn't necessarily appreciate fully. By bringing a few of us together, they helped us develop our vision. They played a catalytic role. They observed what was going on and made a good judgment about what it meant. Once we began discussing it, we began to see what we could do collectively."

In the years since the launch of the multinational project, the *Arabidopsis* research community has become a worldwide network of organizations and individuals. Their continued willingness to share information helps keep the project energized and



An *in situ* hybridization process reveals the accumulation of a specific type of RNA—Apetala 1 —in a mutant Arabidopsis flower. Looking at the flower in this state provides scientists with details about the development of the flower's sepals, petals, stamens, and carpels.

the path cleared for new discoveries. With funding from NSF and other federal agencies, as well as governments in other countries, biological resource centers have been established around the world to make seeds of mutant strains—one scientist called them "starter kits"—available to laboratories that want to study them. Between 1992 and the summer of 2000, the *Arabidopsis* Biological Resource Center at Ohio State University, which shares responsibility with a British center for requests throughout the world, shipped 299,000 seed samples and 94,000 DNA samples. In the spirit of openness and collaboration encouraged by a multinational steering committee, hundreds of *Arabidopsis* researchers worldwide regularly make deposits of new seed lines and DNA libraries into the centers.

Accelerating the Pace

The U.S. component of the multinational effort to sequence the *Arabidopsis* genome started as a joint program by NSF, the U.S. Department of Agriculture (USDA), and the Department of Energy (DOE). Building on this effort, in May 1997 the White House Office of Science and Technology Policy (OSTP) established the National Plant Genome Initiative (NPGI) program with the long-term objective to understand the genetic structure and functions in plants important to agriculture, the environment, energy, and health. In the NPGI's first year, NSF, USDA, and DOE provided additional funds to accelerate completion of *Arabidopsis* sequencing. The international *Arabidopsis* community now expects to publish the complete sequence of the plant's genome by the end of 2000, four years ahead of the original schedule.

In 1999 U.S. and European scientists completed mapping the DNA sequences of two of the five chromosomes of *Arabidopsis* and published their findings in the December 16, 1999 issue of *Nature*. The results—the first complete DNA sequence of a plant chromosome—provided new information about chromosome structure, evolution, intracellular signaling, and disease resistance in plants.

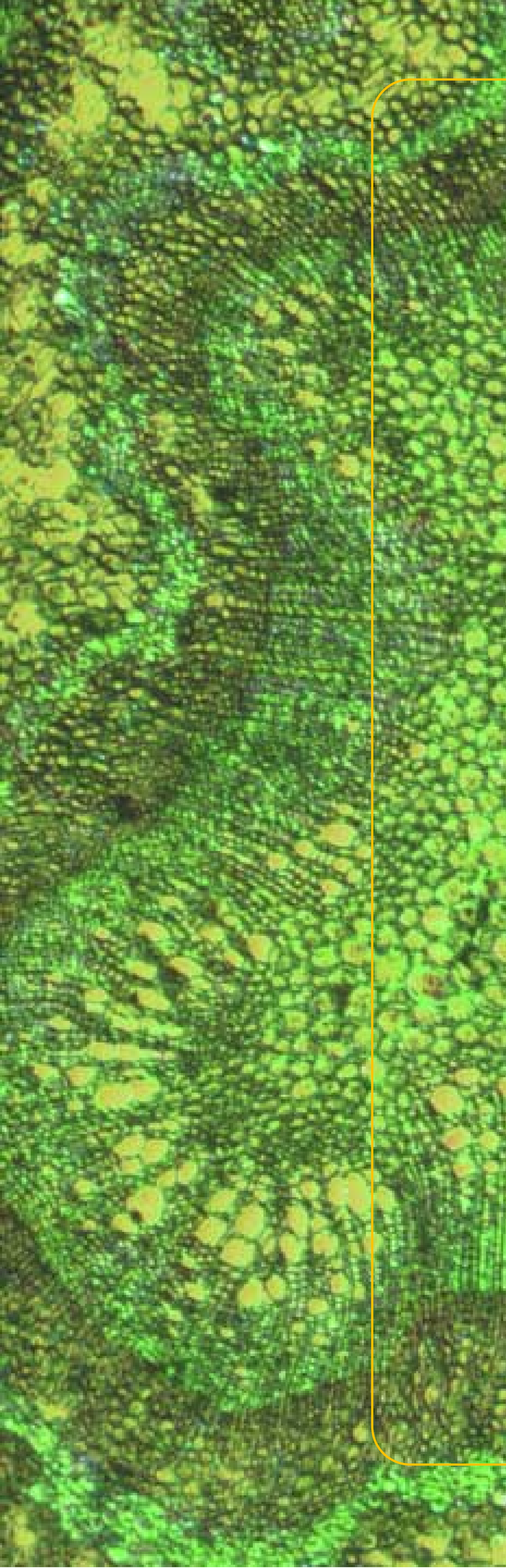
Through the NPGI program, NSF, USDA, and DOE also began jointly funding research to sequence the rice genome, enabling U.S. participation in an international collaboration whose goals are set by the International Rice Genome Sequencing Working Group. Most of the world's major food crops (including rice) are grasses, and they share common sets of genes. The relatively small size of the rice genome—430 million base pairs of DNA divided into 12 chromosomes—makes it a model system for understanding the genomic sequences of other major grass crops including corn, wheat, rye, barley, sorghum, sugar cane, and millet. The working group estimates that researchers could complete the sequencing of the rice genome by 2008.

With rice and other plant sequencing efforts underway and with the completion of the *Arabidopsis* genome sequence tantalizingly close, plant researchers have begun to shift their focus from gene identification to functional genomics—a multidisciplinary approach to develop an understanding of the functions of the plant's genes and how they work together under different conditions. A systematic effort to effectively use the massive amounts of genome data becoming available to determine the functions of all of the genes of *Arabidopsis* is seen as the next frontier in plant research. Such an effort could be accomplished by 2010, according to a recent estimate, and would lead to an integrated database that would be a blueprint of *Arabidopsis* through its entire life cycle.

Research to understand the functions of Arabidopsis's gene sequences are still in the early stages but breakthroughs with considerable practical applications could come in the areas of:

DISEASE RESISTANCE. Plant breeders have long known that certain varieties of crops are more resistant than others to particular viral, bacterial, or fungal pathogens. Disease resistance is a major goal of most plant-breeding programs, but it has typically been a long process involving crop plants found in natural wild populations. The process has been impeded by the "species barrier," which, until recently, prevented desirable genes from being passed around—from corn to cauliflower, for example. *Arabidopsis* researchers have determined the molecular sequences of genes that code for disease resistance and, in addition, the processes by which *Arabidopsis* and perhaps other plants marshal their defenses against pathogens. This discovery may be particularly useful in triggering resistance to disease in species other than *Arabidopsis*. In one enticing example, a bacterial pathogen of mammals was also discovered to be an *Arabidopsis* pathogen. Some of the same factors are required for infection, leading researchers to speculate that evolutionary susceptibility to disease may be accompanied by factors that confer resistance.

CIRCADIAN RHYTHMS. A vast array of processes in plants are regulated in a circadian manner, including daily leaf movements and pore openings, flower-blooming schedules, and photosynthesis cycles. The term "circadian" comes from the Latin words *circa*, meaning about, and *diem*, meaning day. It refers to processes that occur approximately once every twenty-four hours, in response to an organism's internal clock. Plants gain a strong adaptive advantage by being able to anticipate oncoming dawn or dusk, rather than merely responding to the presence or absence of light. They display this ability even as



Communication. . . Fused with the Ideas and Results of Others

NSF's support of genetics dates back to the earliest days of the agency. One of NSF's first five grants in the field of genetic biology, as it was originally called, was made in 1952 to Max Delbrück, who came to the United States from Germany in 1937. Trained as a quantum physicist, he gravitated to biology. While working at the California Institute of Technology and at Vanderbilt University, Delbrück organized and inspired a distinguished group of biologists. One member of the group was James Watson, who, along with Francis Crick and Maurice Wilkins, received the Nobel Prize in 1962 for discovering the structure of deoxyribonucleic acid, or DNA.

Delbrück's contributions to the history of genetics were numerous and evolved from his early interest in bacteriophages, viruses that infect bacteria. A phage can attach itself to a bacterial cell, shuck off its own protein coat, and infiltrate the host cell the way the contents of a syringe enter a vein. Once the phage is inside a cell, its genetic material combines with that of the bacteria, and the phage reproduces itself exactly. These characteristics make phages ideal for the study of biological self-replication and the transfer of bacterial genes between host organisms. Through experiments with phages, Delbrück and a collaborator demonstrated, for the first time, that bacteria undergo mutation. Their work validated the revolutionary idea that genetic principles apply to microorganisms. It also opened the door to genetic analysis of recombination within bacteria.

Delbrück won the Nobel Prize in 1969, as "the man who transformed bacteriophage research from vague empiricism to an exact science." In his acceptance speech, Delbrück remarked upon the ways in which one scientific discovery leads to another, and he contrasted progress in art with progress in science.

"The books of the great scientists are gathering dust on the shelves of learned libraries. And rightly so. The scientist addresses an infinitesimal audience of fellow composers. His message is not devoid of universality but its universality is disembodied and anonymous. While the artist's communication is linked forever with its original form, that of the scientist is modified, amplified, fused with the ideas and results of others, and melts into the stream of knowledge and ideas which forms our culture."

The major multinational effort to understand the intricacies of Arabidopsis has already led to major breakthroughs in engineering disease-resistant plants.

days grow shorter in the fall or longer in the spring. By fusing *Arabidopsis* genetic material with bioluminescent material from fireflies, researchers have been able to observe a glowing pattern of response that reflects the plants' internal clocks. This enabled them to find mutants with aberrant responses, which in turn led to identification of a biological clock gene named "toc." Although influenced by sunlight, "toc" also operates independently, even when the plant is in constant darkness.

ENVIRONMENTAL RESPONSE. Plants respond to a great deal of information from the type of daylight they receive. For example, the changing light throughout the year provides clues about whether it is time to sprout or time to make seeds. When an object blocks the light, plants respond by growing around the object to reach the light. Much of our current information about how plants perceive and respond to light is derived from studies with *Arabidopsis.* These studies have identified the basic genetic framework of light perception and the complex communications system, called a signal transduction network, through which plants act upon information from their photoreceptors. There has also been significant progress in understanding how plants respond genetically when exposed to stresses in the environment, such as ozone, UV-irradiation, touch, cold, and oxygen deprivation.

PLANT HORMONE RESPONSE. Hormones play a central role in the regulation of plant growth and development. Of particular interest is the fruit-ripening hormone ethylene; growers have long searched for a way to minimize crop spoilage by preventing or delaying ripening in a reversible manner. Studies of *Arabidopsis* have demonstrated for the first time the mechanisms through which the tissues and cells of plants respond to ethylene. A gene that prevents response to the growth substance ethylene turns out to be comparable to a "never-ripe" gene in tomatoes, a finding that further supports *Arabidopsis*'s ability to serve as a model for other plants, including crop species.

COMMERCIAL APPLICATIONS. Genetic comparisons between *Arabidopsis* and crop species are increasing constantly. For example, even though the flowers of *Arabidopsis* are very different from those of snapdragons, the same genes control flower development in both. This discovery brings scientists closer to understanding and being able to manipulate the development of grains, fruits, and other flower products to one day create more productive crops. The genes that guide the synthesis of oils in *Arabidopsis* are closely related to those that produce oils in commercial oil crops, a relationship that is already being exploited commercially to produce plants with oils lower in polyunsaturated fats. *Arabidopsis* has also been the test organism for efforts to produce biodegradable plastics in crop plants. Several large chemical companies have started active research programs based on *Arabidopsis* research to develop transgenic crops that produce polyhydroxybutyrate (PHB), a biodegradable plastic.

Why Learn about *Arabidopsis*?

The Multinational Coordinated *Arabidopsis thaliana* Genome Research Project began as an effort by NSF and leading academic researchers to advance fundamental knowledge about how plants function. When NSF published the multinational committee's long-range plan in 1990, U.S. government expenditures on *Arabidopsis* research totaled $7.5 million. In 1993, total expenditures on *Arabidopsis* research by NSF, USDA, DOE, and the National Institutes of Health (NIH) were $22 million. In 1998, NSF, USDA, and DOE awarded an additional $28.3 million over a three-year period to accelerate the pace of *Arabidopsis* research. The global effort to understand *Arabidopsis* involves scientists in more than thirty countries.

From the beginning, NSF saw the *Arabidopsis* effort as an opportunity to foster a collegial, highly motivated, scientific community that would advance fundamental knowledge in an effective way. Both within NSF and in other agencies, officials also recognized that the research has important practical applications. Despite the vast productivity of the agricultural sector, most crops grown in the United States produce less than 50 percent of their genetic potential. Plants succumb to disturbances in their environment; in some years, floods, drought, disease, and parasitic attacks cost billions of dollars. Unlike humans, plants cannot be moved to high ground or inoculated against illnesses. The only protection is to grow resistant strains, and many feel that conventional plant breeding cannot accomplish this fast enough. In the developing world in particular, the problem is exacerbated by growing populations that put extraordinary pressures on the ecosystem. Many see biotechnology as the only feasible solution.

Bioengineered plants also figure prominently in the ideal world envisioned by NIH and the health care community, who see potential in plants as a source of improved, less costly pharmaceuticals. The Department of Energy, for its part, envisions a future in which biotechnology improves the quality and quantity of biomass products, such as alternative fuels and chemical feedstocks, and provides a way to engineer plants to clean up contaminated soil at former nuclear weapons production sites.

Presciently summing up the major applications of biotechnology, a 1995 report from the National Science and Technology Council called *Biotechnology for the 21st Century* stated: “Through the use of advanced tools such as genetic engineering, biotechnology is expected to have a dramatic effect on the world economy over the next decade. Innovations emerging in the food and pharmaceutical sectors offer only a hint of the enormous potential of biotechnology to provide diverse new products, including disease-resistant plants, ‘natural’ pesticides, environmental remediation technologies, biodegradable plastics, novel therapeutic agents, and chemicals and enzymes that will reduce the cost and improve the efficiency of industrial processes. . . [B]iotechnology. . . may well play as pivotal a role in social and industrial advancement over the next ten to twenty years as did physics and chemistry in the post-World War II period.”

To Learn More

NSF Directorate for Biological Sciences
www.nsf.gov/bio

The Multinational Coordinated *Arabidopsis thaliana* Genome Research Project Progress Reports (published by NSF)
www.nsf.gov:80/bio/reports.htm#progress

Gregor Mendel’s work in plant heredity
www.netspace.org/MendelWeb/

***Arabidopsis* Biological Resource Center at Ohio State University**
http://aims.cps.msu.edu/aims/

Nottingham *Arabidopsis* Stock Centre
http://nasc.life.nott.ac.uk/

The Arabidopsis Information Resource (TAIR)
www.arabidopsis.org

National Institutes of Health
www.nih.gov/

National Science and Technology Council
www.whitehouse.gov/WH/EOP/OSTP/NSTC/html/NSTC_Home.html