Yeast, bacteria, and fruit fly geneticists helped bring molecular biology to plant research; now the complete genome sequence of a "model" plant is almost in hand

Arabidopsis Comes of Age

Fresh out of graduate school in 1978, Shauna and Chris Somerville escaped to Paris to figure out what to do with their lives. Both had newly minted degrees in hand; Shauna, a master's in plant breeding, Chris, a Ph.D. in molecular biology, specializing in the lab bacterium *Escherichia coli*. Luxuriating in their temporary jobless state, they spent their mornings at the Institut Pierre and Marie Curie library and their afternoons at various cafés. There, over coffee and croissants, Shauna persuaded her husband to think about applying the tools of molecular biology to plants.

Two papers helped make her case. One, a 1977 report, chronicled the successful use of a plasmid from a soil bacterium, *Agrobacterium tumefaciens*, to ferry genes into plants. The other was a 1975 review extolling the virtues of studying a common, mustardlike weed that was easy to grow in the lab. They decided to wed their two professions and use this weed, *Arabidopsis*, to tackle plant science's more perplexing problems, such as flower development and photorespiration. Their dream, recalls Chris, was to "create a new field," centered on *Arabidopsis*, that would bring molecular biology techniques to bear on plants.

As reports from a September meeting in Miami* make clear, some 20 years later, those café conversations have paid off. The Somervilles—with ample help from researchers ranging from *Drosophila* geneticists to plant physiologists—have firmly established the lowly weed *Arabidopsis* as *the* model organism for plant biology. A remarkably open and collegial community has blossomed around this weed. And, if all goes as planned, by the end of the year *Arabidopsis* will take its place among the pantheon of select research organisms that have had their full genomes deciphered. The genomes of the first three eukaryotes—the yeast *Saccharomyces cerevisiae*, the nematode *Caenorhabditis elegans*, and the fruit fly *Drosophila melanogaster*—have provided glimpses into such fundamental secrets as the signals that control development and the proteins that make up the cell's molecular motors, to name a few. But among these model organisms, only *Arabidopsis* offers the promise of elucidating properties unique to plants, such as how they make seeds and flowers, or how they make efficient use of sunlight. Already, there have been surprises: As described in Miami, plants are more similar, genetically, to humans than to yeast, bacteria, or nematodes. And that's just the beginning, says Athanasios Theologis, a plant biologist at the U.S. Department of Agriculture (USDA) Plant Gene Expression Center at the University of California, Berkeley: "*Arabidopsis* was a tremendous thing to happen in plant biology."

Arabidopsis. Soon to be the first completely sequenced plant.

CREDIT: J. MYLNE/UNIVERSITY OF QUEENSLAND

Scientists in 20 countries, part of a close-knit international consortium known as the Arabidopsis Genome Initiative, are now working around the clock to complete the sequencing by the end of the year. It will be tight, concedes Joseph Ecker, a plant biologist at the Salk Institute for Biological Studies in La Jolla, California, as the consortium is bent on perfection. Indeed, he adds, the first two chromosomes published set a new standard of excellence within the genome community. "There's just a huge dedication to get this as perfect as possible," agrees Rod Wing, a plant molecular biologist at Clemson University in South Carolina.

Modest beginnings

Not even the Somervilles could have envisioned deciphering the weed's entire genome when they returned to the United States later in 1978 eager to begin *Arabidopsis* studies. The field was wide open. True, scattered references to experiments with *Arabidopsis* had appeared in the scientific literature since the late 1800s, and as early as 1943, the German biologist Friedrich Laibach had suggested it might be a useful model genetic organism. A few pioneers had followed that advice, among them George Rédei, a young Hungarian geneticist who brought *Arabidopsis* seeds with him to the University of Missouri, Columbia, in 1957. There he used radiation to create mutants with stunted growth, which he then hoped to use to track down the genes underlying these changes.

Thrilled with the ease of working with *Arabidopsis*, which completes its life cycle in under 6 weeks and grows just a few inches overall, Rédei, too, started writing about the plant's potential. But his colleagues were nonplussed, to say the least. In 1969, "the NSF [National Science Foundation] program director informed me I had to quit *Arabidopsis* if I wanted to continue to get support," recalls Rédei, now retired. But Rédei, who had inherited Nobel laureate Barbara McClintock's old lab on the Columbia campus, fortunately had some of McClintock's stubbornness. "I still continued working [on it]; I knew it was worth doing it."

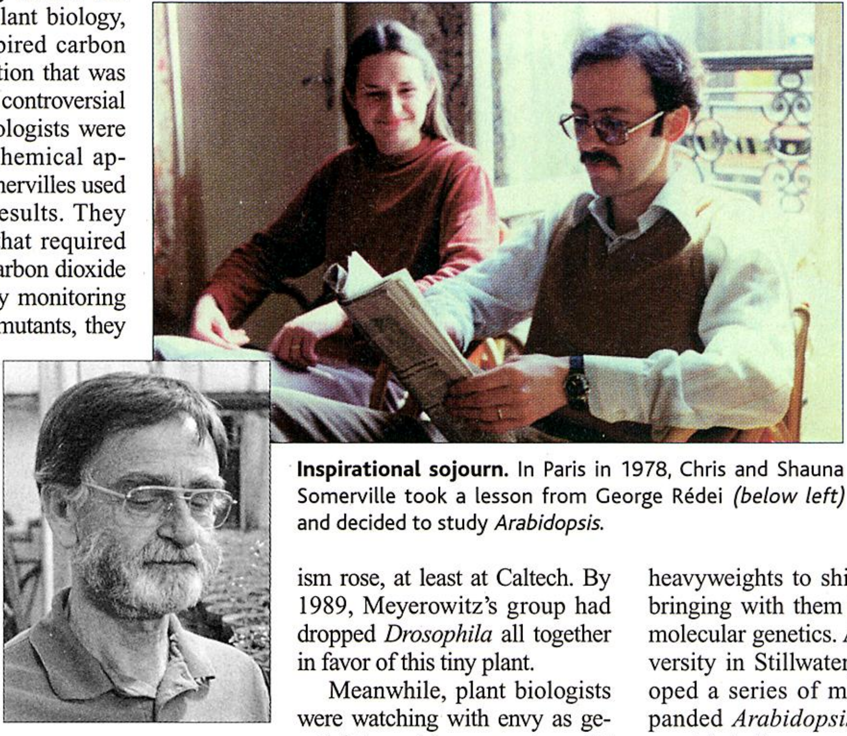
Rédei finally found an ally in Maarten Koornneef, who became interested in *Arabidopsis* as a Ph.D. student at Wageningen Agricultural University in the Netherlands in 1976. Like Rédei, Koornneef began isolating mutants. But he went a step further, using the altered traits and their inheritance patterns to build a detailed genetic map—a key resource for finding specific genes. Even so, for the next several years, Koornneef, his professors, a few other European groups, and Rédei had the field mostly to themselves. They were "the guys that kept the torch burning [for *Arabidopsis*] during the dark ages," notes Robert Pruitt, a molecular geneticist at Purdue University in West Lafayette, Indiana. Indeed, it was Rédei's work that the Somervilles came across during their Parisian hiatus.

---

* The 12th International Genome Sequencing and Analysis Conference, 12–15 September, Miami Beach, Florida.

Chris and Shauna Somerville began their work at the University of Illinois, heading there from Paris to work with William Ogren, a plant physiologist. Fortunately, they came with fellowships that enabled the couple to finance their *Arabidopsis* experiments without having to get grants from NSF.

Never shy, the couple boldly tackled one of the trickier questions in plant biology, namely, where a plant's respired carbon comes from, a research question that was plagued with conflicting and controversial results. While other plant biologists were using physiological or biochemical approaches to sort it out, the Somervilles used genetics, with impressive results. They made *Arabidopsis* mutants that required higher than normal levels of carbon dioxide in the air to survive. Then, by monitoring where carbon built up in the mutants, they identified once and for all the major source of carbon dioxide. With their 1979 *Nature* paper, "we solved the contentious problem of the day," says Chris Somerville. The Somervilles, who are now at the Carnegie Institution of Washington's Department of Plant Biology at Stanford University, then went on to tackle other plant issues—such as the role of plant hormones in regulating growth or responses to infection.



Inspirational sojourn. In Paris in 1978, Chris and Shauna Somerville took a lesson from George Rédei *(below left)* and decided to study *Arabidopsis.*

Grudging respectability

Over the next 5 years, *Arabidopsis*—and the Somervilles' work—began to catch the attention of other researchers, many of whom were geneticists studying other organisms. One was Elliot Meyerowitz, a *Drosophila* molecular biologist at the California Institute of Technology (Caltech) in Pasadena. Curious about plant biology, Meyerowitz was intrigued when one of his students, Pruitt, said he wanted to work on *Arabidopsis*. "In the beginning, [Meyerowitz] said 'Only work on it on the weekends,'" Pruitt recalls. "He saw it more as a distraction from what I was supposed to be doing." But interest within the lab quickly grew, and by 1985, Pruitt and Leslie Leutwiler were working full-time on *Arabidopsis*. In a critical discovery, the pair figured out that the *Arabidopsis* genome was significantly smaller than tobacco's—another favorite of plant biologists. At the time, they estimated that *Arabidopsis* had 70 million to 150 million base pairs (it is now estimated to be 117 million), compared to 1.6 billion in tobacco; what's more, they found it had relatively little repetitive DNA, the bane of gene hunters.

CREDITS: (TOP TO BOTTOM) M. KOORNNEEF; D. DAILEY; THE INSTITUTE FOR GENOMIC RESEARCH

The next year, Caren Chang in Meyerowitz's group cloned the first *Arabidopsis* gene. "We found out that plant DNA was just like any other DNA," Meyerowitz recalls, and thus the molecular techniques that had been developed for other organisms would also work on plants. With these findings, *Arabidopsis*'s stature as a model organism rose, at least at Caltech. By 1989, Meyerowitz's group had dropped *Drosophila* all together in favor of this tiny plant.

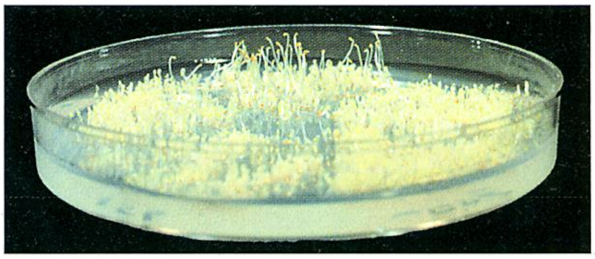
Meanwhile, plant biologists were watching with envy as geneticists working on such model organisms as the fruit fly and the worm reported a series of stunning advances. Complex signaling pathways were worked out in the fly, for instance, while the fate of every cell was traced throughout the life cycle of the nematode. Much of this progress, an increasing number of plant biologists realized, came from a concentrated push on one particular species. At a plant genetics meeting in 1985, "it was clear for the first time that we needed a plant that we could do genetics and molecular biology [on]," Koornneef recalls. But the community needed to pick one.

Over the next few years, while Shauna took on *Arabidopsis* to study plant defenses against pathogens, her husband and Meyerowitz took up the case for *Arabidopsis*, arguing in favor of its benefits over petunia or tomato—two other plants in contention to be a model system. Not only was its genome suitable and its generation time measured in weeks, not months, but *Arabidopsis* also produced thousands of seeds and was self-fertilizing, making genetic studies easier to do. They proved able salesmen and mentors, churning out dozens of young *Arabidopsis* researchers and converting many more. In 1984, just 36 scientific papers dealt with *Arabidopsis*; by 1989 the number shot up to 216 and the number of *Arabidopsis* labs climbed from a half-dozen to more than 100.

A new community began to emerge around *Arabidopsis*. It included not only plant biologists who previously worked on crops but also researchers from labs working on yeast, bacteria, and fruit flies who opted instead to work on this "seed machine," as McClintock liked to call it. Noted yeast geneticists Gerald Fink at the Whitehead Institute for Biomedical Research in Cambridge, Massachusetts, and Ron Davis of Stanford were among the heavyweights to shift into the plant world, bringing with them their expertise in yeast molecular genetics. At Oklahoma State University in Stillwater, David Meinke developed a series of mutants that greatly expanded *Arabidopsis*'s utility for developmental studies.

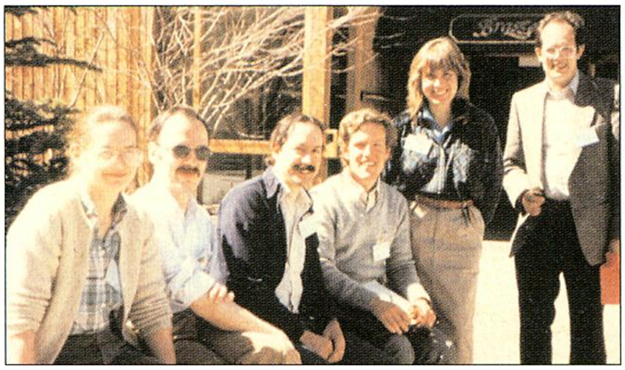
From the outset, the *Arabidopsis* community adopted a culture of openness, sharing data and creating community resources. An international news group, established in the pre-Internet era, "took off like gang-busters," Chris Somerville recalls. By 1992, it had some 1500 subscribers who regularly exchanged information on new mutants and protocols. Two new seed banks were established in 1991—one in Columbus, Ohio, and another in Nottingham, U.K.—relieving the pressure on Albert Kranz, who for decades had been distributing seeds to individual researchers from J. W. Goethe University in Frankfurt, Germany.

Later, in 1993, when Chris Somerville



A model model. Small enough to grow in a petri dish, *Arabidopsis* has many other qualities useful in genetic research.

and colleagues began isolating 30,000 expressed sequence tags, bits of genes that help researchers track down the complete gene, they put the data in GenBank, the public database, before they even looked at them. The effort, funded with a $1 million grant from NSF, would help the entire community, Somerville says, and he didn't think his group should have an unfair advantage. That tradition continues today. Even before the sequencing is complete, the collaborators are developing the next set of joint resources, such as mutants and microarrays necessary for interpreting genomic data. These, along with the genome data, are "leading to the democratization of plant biology," says Jeffrey Dangl, a plant pathologist at the University of North Carolina, Chapel Hill. "Any student in any university with a good idea can access the resources to make discoveries."



Early campaigns. By 1985, a core group of plant biologists was pushing to have *Arabidopsis* adopted as a "model" system. (Left to right) S. Somerville, C. Somerville, E. Meyerowitz, D. Meinke, M. Crouch, M. Koornneef.

The genome era

In the late 1980s, *Arabidopsis* was swept up in the excitement and controversy of the fledgling Human Genome Project. By 1990, the Human Genome Project was well under way, with the express goal of sequencing not only the human genome but also those of several model organisms to enable comparative analyses. (It turns out it is far easier to figure out what a gene does in a fly, say, than in a person—and both, fortunately, share many of the same genes.) James Watson, who then ran the Human Genome Center at the National Institutes of Health, thought that *Arabidopsis* should be sequenced—but he didn't think NIH should foot the bill. So he turned to Mary Clutter of NSF. Intrigued, the agency sponsored a series of workshops that in 1990 resulted in a long-range plan for an international *Arabidopsis* genome project. The goal was to determine the function of every gene.

Sequencing chiefs. A large consortium has been sequencing the *Arabidopsis* genome: (standing from left to right) K. Okada, S. Tabata, A. Theologis, R. Wilson, J. Ecker, S. Rounsley, W. R. McCombie, I. Bancroft, M. Bevan, M. Cherry; (seated from left to right) F. Quetier, D. Meinke, C. Somerville, N. Federspiel, R. Martienssen.

But how that should be done proved contentious. Some, like Meyerowitz, wanted to take a functional approach and track down each gene through mutant screens and classical genetics approaches. Others wanted to follow in the footsteps of the Human Genome Project and sequence the entire genome, predict where the genes were, and then analyze their functions. There were concerns, too, that sequencing would be too taxing for plant science's meager budgets. The functional approach won out for the time being. In its 1991 budget, NSF asked for an additional $5 million to begin those studies. Also that year, the European Commission gave 3 million ECUs to 29 *Arabidopsis* labs. Their goal was to build another type of map of the genome, a physical map consisting of pieces of DNA, and develop new methods for identifying genes. "For [the] first 6 years, genome sequencing was not part of [the program]," says Salk's Ecker.

But enthusiasm for sequencing was growing. Over beers one evening at Cold Spring Harbor Laboratory (CSHL) in New York, Ecker, along with *Arabidopsis* experts Rob Martienssen of CSHL, Joanne Chory of Salk, USDA's Theologis, and Meyerowitz, decided to push for all-out sequencing. They took their case to Watson, who helped them set up meetings the following year that brought together the funding agencies, sequencing experts, and *Arabidopsis* researchers. "The [1994] meetings were very crucial," says Theologis. At one, he recalls, yeast biologist Mark Johnston of Washington University in St. Louis told the group that sequencing was the only way to find *all* the genes; classical genetics just wouldn't cut it. They also heard that sequencing was getting faster and cheaper. "That convinced everyone that [the sequencing] had to be done," Theologis recalls.

A European consortium got off the ground first, funded by the European Union. Building on their experience in mapping chromosomes 4 and 5, 10 labs joined forces in 1995 to sequence chromosome 4. In the United States, NSF, USDA, and the Department of Energy set up a sequencing program that awarded its first grants in 1996. Also that year, the international Arabidopsis Genome Initiative was formed, providing a formal name for the informal collaboration that has worked so well over the years. The group meets yearly and talks far more often to keep the effort running smoothly.

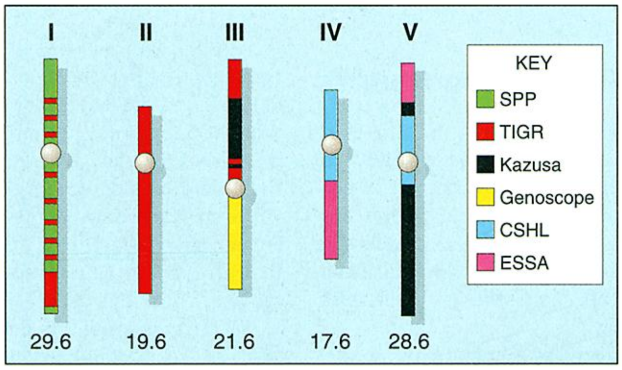
By late 1999, the European Arabidopsis Genome Sequencing Consortium, with a little help from U.S. labs, finished chromosome 4. Right with them was The Institute for Genomic Research, which sequenced chromosome 2 in record time, thanks to assistance from Celera, a new genomics company down the road in Rockville, Maryland. When the two chromosomes were published back to back in the 16 December 1999 issue of *Nature*, they raised the bar for sequencing excellence. In contrast to, say, the *Drosophila* genome, which has lots of missing bits of sequence, or the human genome, in which just gene-rich parts of chromosomes are sequenced, the *Arabidopsis* group sequenced both entire chromosomes, including some of the notoriously tricky centromere region, which most sequencers have purposefully avoided in other organisms because of the low gene content and difficulty. "We did *Arabidopsis* right," says W. Richard McCombie, who has been sequencing *Arabidopsis* at the Cold Spring Laboratory. "It's very satisfying to biologists. And having that information from a very manipulable organism is exceedingly valuable."

Biology transformed

Getting a first glimpse of those sequence data converted any remaining naysayers. "As soon as we started putting the data out there, people got it," says Chris Somerville. Suddenly, the time needed to actually track down a gene dropped from years to months, even weeks. Moreover, the genome has transformed plant biology, says Mary Lou Guerinot, a plant biologist at Dartmouth College in Hanover, New Hampshire. "People aren't thinking about [finding] one gene at a time; they are thinking about gene families, and if there's a [match to the gene] in

CREDITS: (TOP TO BOTTOM) M. KOORNNEEF; AGI

other organisms," she notes.



Who did what. Colors indicate how the sequencing teams—Stanford/University of Pennsylvania/USDA–University of California, Berkeley (SPP); The Institute for Genomic Research (TIGR); Kazusa DNA Research Institute in Chiba, Japan; Genoscope in Evry, France; Cold Spring Harbor Laboratory/Washington University (CSHL); and the European Union Consortium (ESSA)—divided up the sequencing of *Arabidopsis*'s five chromosomes and 117 megabases.

Already, biologists are concentrating on the 8000 or so newly discovered genes that have no known function or matches in other genomes. North Carolina's Dangl, for example, has identified 10 genes that seem to be activated by infection or insect attack—genes that had eluded discovery by his and a dozen other labs for the past 15 years. Now, thanks to ongoing efforts by Ecker and others to create a mutant strain of *Arabidopsis* for each gene, it may soon be possible to work out the function of each of those genes.

A few biologists are also trying to trace the evolution of plant biochemical pathways by looking for the genes involved. By scanning the existing genomic data, for instance, Anthony Bleecker of the University of Wisconsin, Madison, and his colleagues discovered that an *Arabidopsis* ethylene receptor gene is very similar to one in the cyanobacterium *Synechocystis*, a simple photosynthesizing microbe. Because this microbe is thought to be close kin to the microbe that was incorporated into ancient plant cells and became the chloroplast, the similarity suggests that the *Arabidopsis* gene migrated from the chloroplast genome into the *Arabidopsis* genome and subsequently evolved a new role in controlling plant development. "It gives us an idea about how plants got to be so different [from animals]," Meyerowitz explains. "All the sudden we have insight into how things came to be."

Not bad for a plant that few biologists had heard of a decade ago. Given these and other insights, plant biologists are clamoring for the sequence—but finishing it by the end of the year, when a paper is due at *Nature* and a party is planned at Cold Spring Harbor Lab, is going to be a stretch. "We're close, but we're not done," concedes Michael Bevan, a plant biochemist at the John Innes Centre in Norwich, U.K. Meanwhile, in Japan, the United States, and Europe, computers and sequencing machines are running full tilt, polishing off the last pieces and filling in gaps (see graph at left). At the same time, two of the labs are feverishly scanning the entire genome, predicting and classifying genes. It's a frantic pace, involving scores of people—quite a contrast to the life the Somervilles were enjoying when they first envisioned this plant's future 23 years ago in Paris.

–ELIZABETH PENNISI

ECOSYSTEM RECOVERY

Ghost Towns Tell Tales of Ecological Boom and Bust

Scarred desert ecosystems are recovering at the sites of some abandoned boomtowns, but are slow to heal at others. Soil age appears to be the key difference

The greatest mining scandal in U.S. history struck Greenwater, California, in 1906. A miner discovered copper in Death Valley's Black Mountains—and within months, this boomtown exploded with saloons, shacks, and 2000 treasure hunters. "Greenwater is destined to be the richest mineral-producing city on the whole globe," declared one flyer.

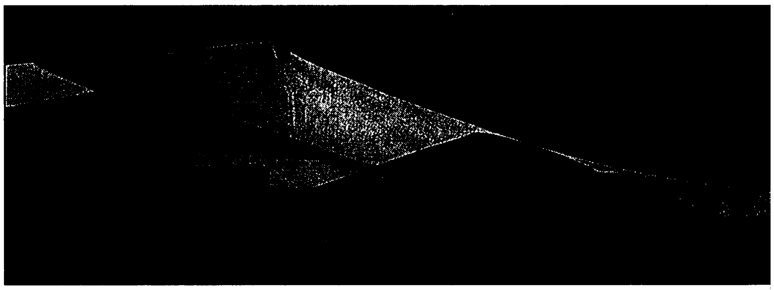
By 1908, the town was empty.

Greenwater turned out to be more bust than boom. Only flecks of low-grade copper graced the mountains, while swindlers made a fortune selling shares in nonexistent mines. But today, Greenwater is finally paying off—with ecological data. Robert Webb, a hydrologist with the U.S. Geological Survey (USGS) in Tucson, Arizona, is studying eight Mojave Desert ghost towns, including Greenwater, to see how their long-abandoned soils have recovered from the pounding they took during the mining boom. Sprawling creosote, burrobush, and desert sage are reclaiming these boomtown sites, turning them into outdoor labs for desert recovery—and challenging popular ecology theory along the way.

For at least 25 years, ecologists have described the parched desert landscape as "easily scarred and slowly healed." Conventional wisdom once held that disturbed desert soils never fully recover. Without question, the landmarks are stark—in parts of the central Mojave, for instance, tank tracks left from World War II training exercises still gouge parts of the desert floor. But, popular wisdom aside, Webb and other scientists are also witnessing desert recovery.

In Webb's study, disturbed soils at some ghost town sites have rebounded, defiantly blooming in less than a century—while others sport much sparser plant populations. By comparing the towns' geology and history, Webb is uncovering features, such as soil age, that determine rates of plant recovery. The emerging data could help the managers of the Mojave's six military bases and four national park areas make ecologically sound management decisions. "All landscape is not created equal," Webb says. "If you have to route a road, if you have to run tanks over it, let's use our scientific sense about how to do it."

Resource managers say insight from the Mojave is badly needed. "There's a lot of pressure on park managers to make quick decisions, and we're starving to find immediate answers to problems," says vegetation specialist Jane Rodgers of Joshua Tree National Park, which spans parts of the Mojave and Colorado deserts. "If we're looking at where to put a new campground, for instance, this information could be really useful." The ghost towns also offer a rare look at desert recovery in action, adds Joseph



Fragile beauty. Death Valley is fragile in parts, resilient in others.

CREDIT: (TOP) ADAPTED FROM TIGR, WITH INPUT FROM J. ECKER; C. AURNESS/CORBIS