A DISRUPTIVE PERSONALITY, DISRUPTED

Eric Schadt revels in making people uncomfortable with his science. **Bryn Nelson** reports how the bioinformatics rabble-rouser hopes to charge ahead in the face of his company's disintegration.

In need of an escape from the mental gymnastics of hardcore genome analysis, Eric Schadt, executive scientific director of research genetics for Rosetta

Inpharmatics, is clear about what works for him — careering down a steep mountainside on a snowboard. “You can’t sort of ease your way down a hill,” he says over breakfast near the company’s headquarters in Seattle, Washington. “It’s either all, or nothing.”

That fearless approach may be tested after a bombshell announcement last month. Rosetta, which has been on a head-turning run for most of the past decade, now finds itself in mid-air, hoping to make a landing that could be very tricky indeed. Merck & Co., which bought the biotech company in 2001 and operated it as a subsidiary, announced on 22 October that it will close down most of the Seattle operations by December 2009, transferring Schadt and a number of his team to its Boston research centre.

The move comes as part of a global reorganization in the face of slumping drug sales, and includes cutting more than 7,200 jobs worldwide. It could be worse. Reni Benjamin, a senior biotechnology analyst and managing director for investment bank Rodman & Renshaw in New York, says that given the current economic climate for pharmaceuticals, sub-

sidaries such as Rosetta have no guarantee of survival. “They could just cut the entire thing and call it a day if they didn’t think it was important.”

What’s likely to save Rosetta from complete oblivion is Schadt’s trend-setting science of integrated genomics, which uncovers disease mechanisms by revealing vast networks of gene interactions. Genome-wide association studies, which have been a favoured technology for finding disease-linked genes over the past several years, seek out associations between a disease state and a genetic variant. Although the studies have turned up hundreds of variants associated with disease, they can detect only independent effects of individual genes, which means they might miss a lot (see page 18).

Like genome-wide association studies, Schadt’s approach tries to correlate variations in DNA with some observable complex trait, such as a disease, in a population. But Schadt and his group add a third factor: gene-expression levels, as measured by microarrays. They then use the data to build models of how the three factors — variations in the genotype, variations in the disease, and variations in gene expression — are related.

Some relations are straightforward: a gene variant has a direct effect on expression, and that has a direct effect on the disease trait. But it is also possible for a particular genetic variation to be linked to a disease trait that, in turn, alters the expression of some other gene. And then there are cases in which a genetic variation is linked to both a disease and an unrelated change in gene expression. The complexity of all this goes through the roof when genes start interacting: the models explode into networks of interconnected elements. But these network models allows

Schadt and his colleagues to identify and validate associations between genes and disease — and to predict how perturbing the system, with a drug or genetic mutation, will affect expression and disease.

Success story?

So far, company officials say, the strategy has worked well.

Of the 52 compounds in Merck’s clinical-trial pipeline in 2006, ten entered through Rosetta’s efforts. Now, according to Stephen Friend, a senior vice-president at Merck and Rosetta’s co-founder, the approach is so integrated, it’s hard to tell what originated from Schadt’s team. Something that important

“Their whole lives were ‘What the heck is that gene?’; my point was ‘Most of those genes aren’t even druggable’.”

— Eric Schadt

is not going to be axed just to control costs, says Benjamin. "If 20 to 25% of a company's pipeline is being generated from one particular platform, you would have more significant clues than 'streamlining operations' if they didn't like what they were seeing."

On Rosetta's fourth floor, there's all sorts of stuff that may not survive the move east. A cheeky monument built from lab equipment including discarded flasks, racks and pipettes, entitled 'Don't Mess With Us, We're Scientists', sits near display cases showing off the company's early innovations such as its microarray spotter prototypes. One recent acquisition is likely to stay on the packing list. The Hamilton MicroLab Star, a custom-designed platform hosting three interconnected robots, represents the next generation for gene-expression profiling, offering a nearly seamless start-to-finish automated process.

Upstairs, John Dey, the company's UNIX operations manager, lifts a floor panel to reveal a sea of blue cables and a rough visual gauge of the computing power housed within the company's nerve centre. The cluster began with eight computer nodes in 1998. Now it boasts 1,000 connected by more than 16 kilometres of cable. The computational muscle in Boston will have to be bigger still. Dey shakes his head when asked whether he's confident about keeping up with storage demands. "Oh, it's going to explode," he says.

Schadt's team needs this state-of-the-art equipment for what Schadt calls one of the biggest looming bottlenecks for biology in the next ten years: understanding the effects of many genes interacting at once. For genome-wide association studies, the question is fairly straightforward from a computational perspective. But testing biology's dizzying network of interactions in a holistic way, Schadt says, requires the computational power generally reserved for climate modelling and astrophysics.

This aggressive approach rightfully has made specialists uneasy. Friend says integrated genomics met with an initial 'show me'-style distrust. With any new technology, he says, an evangelical wave first oversells it and turns almost everyone into a sceptic.

Schadt clearly revels in proving the critics wrong. In May, Rosetta led a study of gene expression in the human liver and found more than 6,000 associations between single nucleotide polymorphism genotypes and gene expression traits¹. Although many of the expressed genes were already implicated in

human disease, other connections were newly exposed as suspect. The study's integrated data, for example, pointed to *RPS26* instead of the previously favoured *ERBB3* as the most likely susceptibility gene in a genomic region associated with type 1 diabetes.

The layering approach by which Rosetta constructs a complex network has garnered its own share of detractors. "To some of these naysayers, you have a big magic black box where you pour in everything and out comes a drug target, and that doesn't sound like science," says Trey Ideker, an associate professor of bioengineering at the University of California, San Diego, who is collaborating with Schadt on a review detailing the use of systems and network modelling for drug development and health care. "But if you look under the hood, what Eric is doing is absolutely sound — it's the sheer complexity that is overwhelming."

Friends reunited

Schadt often winds his critics up, of course. "The network stuff still makes a lot of people uneasy," he says. "People don't want things to be that complicated." He smiles as he recalls his admittedly "inflammatory" talks a few years ago, in which he basically told other scientists: forget genes and focus on what genetic perturbations are doing to the whole system. "Human geneticists just hated it when I would say stuff like that," he says.

"You know, their whole lives were, 'What the heck is that gene?' and my whole point was,

'Well, first of all, most of those genes aren't even druggable.'"

Most of Rosetta's methods have all but ignored the question of gene identity initially and instead tracked disease-associated hiccups in a genetic network. The approach led to the 2005 publication in *Nature Genetics* of a study

laying out the general integrative strategy — something Schadt counts among his proudest accomplishments². "It wasn't just the intellectual satisfaction, it was that everybody — nearly everybody — was saying, 'Nah, it's not going to work,'" he recalls. He looks back almost wistfully on those earlier fights. "Because our work has got so successful now, I don't feel that

people push back enough," he says. "It's almost too accepting."

The troubled economic landscape could provide plenty of pushback to Schadt's resource-intensive approach, but he seems unfazed by unanticipated changes. Unplanned course corrections have defined his past. Raised in rural Michigan by a conservative family that placed little value on education, Schadt joined the US Air Force after high school and enrolled in its elite pararescue programme, sometimes called 'superman school' because of its gruelling training regimen. During one exercise, Schadt dislocated his shoulder so badly he required surgery, ending his rescue career. His superiors, though, noticed his high test scores and asked about college. It had never occurred to him.

With financial assistance from the military, he gravitated towards mathematics and computer science at California Polytechnic State University in San Luis Obispo. "It just opened up a brand new world." It wasn't until he earned

"What Eric is doing is absolutely sound; it's the sheer complexity that is overwhelming."

— Trey Ideker

A Merck scientist prepares samples for hybridization to DNA microarrays.

his master's degree in pure mathematics from the University of California, Davis, that biology first caught his eye. Schadt began attending biology seminars and met Kenneth Lange, now chairman of the human-genetics department at the University of California, Los Angeles. Lange encouraged Schadt to pursue a curriculum in biomathematics and eventually became his PhD adviser, guiding him to a degree that integrated molecular biology and applied mathematics. "It just all made sense," Schadt says.

By the time Merck began courting Rosetta in 2001, Schadt had become a key member of the team behind a seminal annotation of the human genome³. Annotation aided by gene-expression microarrays was fast gaining attention, Schadt says, "and we just thought we were going to make it big".

The company, in fact, was on the verge of agreeing to a lucrative contract to expand on its annotation work. "The day we were supposed to sign that deal, Stephen Friend comes in and says, 'We're not going to sign,'" Schadt recalls. "So I said, 'Do you have a better \$20-million deal?'" Merck ended up buying Rosetta in a stock swap valued at about \$620 million.

Next year, Schadt and Friend will be reunited in Boston, where Merck is hoping for a better synergy by co-locating its molecular-profiling and integrative-genomics research with Friend's oncology group, as well as the researchers working on bone and respiratory diseases, arthritis and analgesia. "By providing a more comprehensive view of the numbers of genes that may be causally related to disease," Schadt says, "we can actually use the networks to prioritize those targets."

Identifying and halting a drug programme headed for trouble can be just as important, and Schadt cites the *Ppm1l* gene as a perfect example. In March, his team found an obesity connection for *Ppm1l* — which encodes a newly discovered but poorly characterized protein phosphatase — and two other previously unlinked genes. But the group found that *Ppm1l* also modulated traits for diabetes and high blood pressure⁴. "What we showed is when you knock the gene down, you improve your diabetes profile or you become less insulin resistant, but you also gain weight." Even worse, knocking down the gene also increased blood pressure. Making someone heavier and with higher blood pressure in exchange for a lower diabetes risk, the company concluded, wasn't worth the trade-off

— something it might have missed by taking a narrower focus.

Thomas Gingeras, head of functional genomics at Cold Spring Harbor Laboratory in New York, says Schadt should be commended for embracing a systematic approach to teasing out the molecular networks in cells. But he worries about the initial downside to such an ambitious plan. "The concern I have is whether the information we have to be able to put together this network is patchy and, in some cases, unreliable; and, even if it is accurate, whether we have the right interpretation of those data," he says. By focusing on the low-hanging fruit within the network, might researchers be losing sight of more important non-coding elements that aren't yet within reach?

A legacy lives on

Schadt is confident that his team's models can adapt as a wider range of information comes online, including forthcoming studies that will incorporate data from genome-wide screens of small-molecule metabolites and protein-protein interactions. He's particularly enthusiastic about several big pilot projects that are resequencing entire transcriptomes for hundreds of mice and humans. They offer a way to ask whether largely unknown non-coding RNAs may be central players in the protein-coding network.

Ritsert Jansen, a bioinformatics expert at the University of Groningen in the Netherlands, says Schadt's ability to use high-throughput screens and work with large patient

populations is a big step closer to explaining the causality of complex traits. Jansen, who works on molecular networks in the *Arabidopsis* plant model, says Schadt's work has so far suggested that a few very influential DNA variations are critical for linking genotype with phenotype. The expanding repertoire of 'omics-based studies, including epigenomics, should lead to more dynamic models that zero in on the most important perturbations, he says.

One of the next phases in that progression will add in more clinical information from Merck collaborators such as the Moffitt Cancer Center and Research Institute in Tampa, Florida. Every cancer patient giving informed consent will have multiple tissue samples collected and analysed, with an eye towards charting tissue-to-tissue communication networks.

And then what? Rosetta's legacy may be in providing Merck with a mastery of biological information and superior predictive models of disease risk, drug development and patient response. But ultimately, Schadt says, consumer genome-testing products, such as those provided by 23andMe in Mountain View, California, Navigenics in Redwood Shores, California, and deCODEme of Reykjavik, Iceland, will take the lead in solving perhaps the problem of the century: translating all that information for the people who need to know what it means, be they doctors or parents. "The next ten years are going to be an amazing ride to see how this all plays out," he says.

Despite Schadt's sentiment that "change is good from time to time", he concedes that keeping his team together and focused on its research during the move east will be challenging. But his overall mission, he says, remains unchanged. In an e-mail to *Nature* shortly after Merck announced it would close Rosetta and two other research units in Japan and Italy, Schadt wrote: "It is very gratifying that our work is so highly valued within Merck and that as a result it will become even more integrated within Merck's research." The bottom line: "everything is continuing, we will just be doing it in Boston instead of Seattle". ■

Bryn Nelson is a freelance science writer based in Seattle, Washington.

1. Schadt, E. E. *et al.* *PLoS Biol.* **6**, e107 (2008).
2. Schadt, E. E. *et al.* *Nature Genet.* **37**, 710–717 (2005).
3. Shoemaker, D. D. *et al.* *Nature* **409**, 922–927 (2001).
4. Chen, Y. *et al.* *Nature* **452**, 429–435 (2008).

See Editorial, page 1, and News Feature, page 18.

"It wasn't just the intellectual satisfaction, it was that everybody was saying, 'Nah, it's not going to work'."

— Eric Schadt

Rosetta's 1997 technology for studying yeast expression was quickly replaced by newer tools.