

Hunting the Hot Spots

When Evan Eichler started noticing genetic duplications, he brushed them off as anomalies.

Now the leading expert on the subject, colleagues say his research is one of the best things to come out of the genome.

By Meredith W. Salisbury

As with so many things in life, this story begins with a pair of rabbits. Evan Eichler was raised on a farm in remote Minitonas, Manitoba, population 500, some 500 miles north of the Montana border. When he was about 10 years old, his mother decided she wanted some angora sweaters. Rather than opening up the Sears catalog, they ordered a pair of angora rabbits, and Eichler got the job of breeding them to produce the right colors so she wouldn't have to dye the homespun wool. He wound up with a colony of 200 rabbits — and quite an interest in genetics.

For Eichler, now 33, an avid genealogist and the director of the bioinformatics core facility at the genetics department of Case Western Reserve University, that interest never died. It turned out to be good news for all the

genomics researchers who now rely heavily on his extensive work with duplications in genome sequence.

COMING CLEAN

After getting his biology degree in Canada, Eichler set his sights on a graduate program at the University of Toronto or Johns Hopkins, but underdog Baylor responded to his application with unexpectedly heavy recruiting. Intrigued, he studied up on the school. "I was just so impressed by what these people had done." So he packed up his black Plymouth Turismo and made the trek to Texas.

He joined David Nelson's lab, where the staff had just discovered the gene for fragile X syndrome. On a lark, Eichler began mapping the tip of the X chromosome when something unexpected happened: "When we were mapping some

of the probes they went to different locations in the genome," he says of the genetic material, including a whole gene, that was found on other chromosomes. "We thought chromosomal regions were going to be kind of sacrosanct. ... The whole idea of transposition ... hadn't really been discussed or considered for the human genome."

Though Eichler was fascinated by the finding of a new kind of variation, most people chalked it up as an anomaly, and he turned back to studying fragile X repeats. He went on to postdoc at Lawrence Livermore National Laboratory in 1995, looking into gene duplications on the side but concentrating on other problems.

By 1997, he was ready for a faculty post. He was drawn to the young, enthusiastic staff at Case Western. "My wife was like, 'Oh, no, Cleveland — we're not moving there,'" he remembers. But Cleveland it was, and the move enabled him to shift his attention to the duplications he'd been itching to study. "It was the first time I came clean with the rest of the scientific community and said, 'This is really what I want to work on' — the process or mechanism, the importance of which hadn't been validated or justified yet."

That was a prescient choice. Richard Gibbs, director of the genome sequencing center at Baylor University, remembers the void in duplication research right around the time Eichler moved to Case. Five years ago, Gibbs was looking for "someone to write a good review



Photo: Roger Mastroranni

about what's been learned about structures of duplications and their role in the genome" for publication in *Genome Research*. "There was not really someone we could identify readily who had done actual work in the area," Gibbs says. "That's about when Evan started ... he really filled a niche."

RACE TO PUBLISH

Since that time, Eichler has made a name for himself as the maestro of duplicated regions. The real seal of approval came in late 2000: a personal invitation from Francis Collins to join the Human Genome Project.

Eichler readily agreed to sign on for what he calls "the paper of a lifetime," but his dozen-strong lab wasn't quite ready. "I hadn't thought about how we'd do this on a whole-genome scale," he says. "We pretty much had to reinvent ourselves into kind of a bioinformatics group." That meant begging his department chair for money for a computer cluster — it started out as 35 nodes with half a terabyte of storage on a \$30,000 budget.

With just "seven months to pull this off," Eichler and his crew worked furiously to get the cluster running, only

to have it crash, data unsaved, every time they got it up. Eventually it hit them: the disks were spinning so fast that the RAID was overheating. Eichler likes to joke that after all the work and equipment that went into the cluster, "what we really needed were two \$15 Kmart fans to blow on this thing constantly."

Jeff Bailey, a student in Eichler's lab, recalls the mad dash toward getting into the genome project. "Algorithms [for this scale] hadn't really been considered," he says. "People were just throwing stuff together."

CLEANUP CREW

Eichler, now a father of three, continues to push the envelope in his field as he studies these genome hot spots. One gene his team identified was dubbed "Morpheus" because it's changing so fast — 50 to 100 times faster than most genes — that no one can figure out its function. These regions are implicated in evolutionary divergence, such as between chimp and human, or orangutan and great ape. "The genome ... has a lot more of this duplication than we thought," he says. "It occurs at

nonrandom places, it occurs at specific time points."

That there's a pattern to the duplications has given Eichler a potentially powerful view of the genome. "We now can identify the regions of the genome that are going to be susceptible to gains and losses" — and therefore, to disease-causing deletions or copies of genes.

His niche has afforded him an enviable vantage point. Eichler got access to do side-by-side comparisons of public data with Celera's random read data. "The public side and the private side did a really bad job in this type of region," he sums up. "They're underrepresented, misassembled, mismapped."

Mark Adams, VP of bioinformatics for Celera who has known Eichler since his pre-Celera days at TIGR, helped with the project, which was scheduled for publication in *Science* last month. "He's a very even-handed, creative thinker," says Adams. "That's why we've been able to get along with him very well, however he moves in the political circles of the genome [community]." (Eichler is the first to admit he sometimes has to be pulled off his soapbox, which he mainly uses to admonish the public side to finish the last complicated regions of the sequence.)

"It's one of my favorite genome stories," Adams says of Eichler's work. "He identified a gene family in the duplicated region that appears to be under positive selection. ... The exons are changing faster than the introns. It's just absolutely stunning."

By the nature of his work, Eichler's team is involved in the most intractable portions of the genome, where it's nearly impossible to tell a missed overlap from a duplication. "Our lab sometimes feels like we're a big cleanup crew going in, cleaning up after all the celebrations are over. It's fundamentally important to understand this property of the genome," he says, edging toward his soapbox.

"The final chapter hasn't been written," he says. "Following [my] hunches on these things turned out to be really rewarding." 