
TWENTY-TWO YEARS AGO, David C. Page went fishing for DNA. Poking randomly at a collection of human DNA fragments in a laboratory gene-mapping experiment, the young medical student happened to pick up a Y-chromosome fragment. “When people ask me how I chose the Y to study,” Page deadpans, “the correct answer is ‘with a toothpick.’” ¶ Chance may have chosen the direction, but rampant curiosity and a measure of stubbornness have since propelled Page’s quest to understand the Y—a strange, shrunken chromosome, once dismissed by many as a genetic wasteland too barren to be worth studying. ¶ Page, an HHMI investigator at the Massachusetts Institute of Technology’s Whitehead Institute for Biomedical Research, has repeatedly proven the skeptics wrong. Now his odyssey has reached its most dramatic point yet: In the June 19, 2003, issue of *Nature*, a 40-person team led by Page reported on the decoding of the DNA sequence of the “male-specific,” or MSY, region that makes up 95 percent of the Y chromosome. The MSY region

DAVID PAGE SAYS THE **Y** CHROMOSOME HAS A FEW SURPRISES UP ITS SLEEVE. **BY RICHARD SALTUS**



PHOTOGRAPH BY KATHLEEN DOOHER



David Page studies the Y—a strange, shrunken chromosome, once dismissed by many as a genetic wasteland too barren to merit much attention.

contains the gene that determines an embryo's sex—and was once thought good for little else.

DIVERSE LANDSCAPE

The MSY sequence was wrested from confusing arrays of 23 million letters of genetic code (nucleotides, or base pairs) in a marathon effort that one commentator, Huntington F. Willard, at Duke University's Institute for Genome Sciences and Policy, termed "heroic." A team of sequencing experts led by Richard K. Wilson at Washington University School of Medicine in St. Louis needed more than two years to sort out the maze of DNA sequence patterns in the MSY region. The effort was handsomely repaid: Page and colleagues have described a diverse landscape on the Y that contains 78 important protein-coding genes—hardly a genetic wasteland—and some surprising chromosomal features. The effort has "helped us rethink and throw off so many old, inaccurate ideas about the Y chromosome all at once," Page says. "We had glimpses of these things over the last six or eight years, but it took the complete sequence to show convincingly that these were not just anecdotes, but that they were the themes."

Each theme is a story in itself: how the Y chromosome and its female partner, the X, were once nearly identical but lost the ability to exchange most of their genes and went separate ways; how the Y dwindled to a fraction of its former size, gaining the reputation of a "rotting chromosome," yet harbors a critical set of genes for sperm production; the finding of an unexpected mechanism that enables the lonely Y to exchange DNA within itself (for the most part, it cannot exchange genes with the X chromosome); and the recognition that males and females differ genetically to a far greater degree than had previously been supposed—something that is certain to stir debate and further research.

Page gives much credit to two colleagues in his lab, computational biologist Helen Skaletsky and pediatric endocrinologist Tomoko Kuroda-Kawaguchi, for their zealous, almost obsessive, attention to detail in deciphering the meaning of the Y sequence. "Most people would either have given up or found shortcuts, and then we wouldn't know what we know," Page says. Another colleague, Steve Rozen, led the work that revealed the Y's odd gene-exchange method.

Most of the MSY sequence was in hand—and on the Internet—as much as three years before the *Nature* publication. Page and his colleagues spent that time reading and rereading the text "to see if we could extract some larger meaning from it." Page had never doubted it would be an amazing tale. "It was clear that the Y had a lot of surprises up its sleeve," he says.

HALL OF MIRRORS

The findings that Page's team described at a June press conference made headlines around the world. With his characteristic flair for metaphor, Page hailed the most spectacular Y-chromosomal feature as a "hall of mirrors" and a "crystal palace." He was referring to immensely long structures containing many repeated sequences and multiple copies of genes—scientifically termed "amplicons." In these regions, the researchers had identified eight complex and precise palindromes—segments along which the DNA code reads the same in one direction as the other.

Although palindromes are found on other chromosomes, Page says, they are not as large or as complex. The Y's palindromic sequences form pairs of symmetrical arms that extend in opposite directions from a central hub, spanning up to nearly 3 million nucleotides from tip to tip. Page calls this largest of the Y's eight palindromes "mind-boggling" because it accounts for a full one-thousandth of the entire genome. On six of the palindromes, the researchers identified eight distinct genes, expressed predom-

inantly in the testes, that govern spermatogenesis, the making of sperm.

Some of the palindromes were first discovered several years ago, and they had given genome sequencers fits. DNA is decoded by breaking up the chromosome into small segments, then determining the order of nucleotides on those bits, and finally fitting them back together by matching up distinctive landmarks with the aid of powerful computer algorithms. But here, the long identical stretches offered few nucleotide differences for the scientists to use as landmarks; it was difficult to know even which arm of a palindrome a given piece was on.

"We had a big bin of things we couldn't assemble," says Wilson of Washington University. "It was analogous to a 10,000-piece jigsaw puzzle of ocean and blue sky and a little sailboat. All the pieces look the same. You're hoping you can find a small bird in the sky." For just this reason, the sequencers used DNA from the Y of a single man, an anonymous volunteer from the Buffalo, New York, area. Had the researchers combined DNA samples from several individuals for sequencing, as in the Human Genome Project, normal genetic variation would have washed out the tiny sequence differences the scientists sought in order to get a toehold on the Y. Success came slowly and iteratively, with the finding of differences among a few base pairs, the placement of those landmarks on a map, and then sequencing and mapping—over and over.

Francis S. Collins, director of the National Human Genome Research Institute at the National Institutes of Health, says, "More than any other center, the Washington University group perfected the map-based sequencing approach—and this was absolutely essential."

Once the hurdles were overcome, the hall-of-mirrors structure—with identical sequences on the palindrome's twin arms, capable of facing each other—could be seen as a bulwark against the relentless loss of genes that has been whittling away at the Y chromosome for eons. Such losses have resulted from the Y's inability to recombine, or exchange genes, with the X, as other chromosome pairs do to cull out harmful mutations and maintain their health. To retain at least some functionality, the Y can swap DNA internally: Evolution apparently provided a protective mechanism called "gene conversion," dependent on these extra gene copies and paired palindromes, for repairing mutated genes on the Y.

In the rest of the human genome, the 22 pairs of autosomes (non-sex chromosomes) line up with each other when cells are preparing to divide to create sperm and eggs for the next generation. At this time, the chromosome pairs often cross over each other and exchange chunks containing a few or many genes. In the gene swap, sequences having "typographical" errors involving only one or two genetic letters or large regions of missing or rearranged DNA that cause functional changes (mutations) can be rectified by substitution of normal sequences for the damaged ones.

The paired X chromosomes in female cells recombine this way. But the Y is the odd man out, so to speak, and lacking a trading partner,

Recombination Station

Unlike the two X chromosomes in females, the Y chromosome does not have a partner with which to swap genes in order to replace mutations. The Y appears to protect its genetic integrity by swapping multiple copies of the same gene within its own structure. These graphics suggest how a mutation on a palindrome in the Y (yellow band, figure 1) can be overwritten and thus repaired by a normal gene (2). Sometimes, however, the opposite happens—a mutation overwrites the correct sequence on the other arm and duplicates itself (3). To see these processes in animation, see the links at www.hhmi.org/news/page5a.html.

1



untold numbers of genes have become irreparably mutated and lost over millions of years. (Because the genes that the Y does retain are handed down clonally, not sexually, from father to son to grandson and so on, the Y does not pass through the female line.)

OVERWRITING ERRORS

The process of gene conversion has been most thoroughly studied in yeast, but in these simple organisms, no palindromes are involved. On the Y chromosome, gene conversion occurs when the palindrome's twin arms swivel so that their identical sequences face each other, and if one arm contains a mutated gene, a normal copy on the opposite arm can overwrite the error. (Sometimes the opposite happens, however—the “bad” copy overwrites the correct sequence on the other arm.) Gene conversion occurs on a smaller scale elsewhere in the human genome, “but it’s always been viewed as an event that’s so occasional it’s noteworthy when it happens,” says Page. “What we’re saying now is that in the case of the Y chromosome, it’s standard operating procedure.” In fact, when they traced back the gene configurations on the palindromes of the Y chromosome of both

off from each other. Among these transposed sequences, the researchers could detect only two genes.

Page and his colleagues have constructed from the Y sequence data a history of human chromosome evolution. It begins some 300 million years ago, when our reptilian ancestors existed as males and females but sex determination was probably by some environmental factor, not sex chromosomes. At some point, one of an identical pair of ordinary chromosomes, or autosomes, acquired a mutation that determined maleness. Thus, the Y chromosome, with its sex-determining gene, was born. For a while it continued to recombine genes with its partner (now termed the X). But gradually the two stopped recombining along most of their length, apparently as a result of catastrophic events over millions of years in which large blocks of the Y became inverted, shutting down the DNA-swapping process one stretch at a time along the chromosome. (Two regions, one at each end of the Y, continue to recombine with counterparts on the X chromosome.)

At this point, the Y began to “rot.” Absent recombination, genes mutated and became junk littering the MSY region, and the functional parts of the chromosome withered. “During the last 300 million years,”

THE Y CHROMOSOME HAS THROWN OVERBOARD THE VAST MAJORITY OF ITS GENES.

humans and chimpanzees, the researchers concluded that gene conversion was already present in our common primate ancestors.

As published in a second paper in the same issue of *Nature*, Page and his colleagues calculate that for every boy born in recent times, an average of 600 nucleotides have been swapped between palindrome arms. The effects of most of these exchanges would be unnoticeable. Evidence suggests that gene conversion repairs mutations on the Y as fast as they occur—good news for the chromosome’s continued existence.

The amplicon regions and their grand palindromes make up about one-quarter of the euchromatic (actively expressed) DNA in the MSY region. But in addition, the Page team described two other classes of genes, along with large amounts of heterochromatic (nonfunctional) DNA sequences. Twenty-nine genes on the Y are called “X-degenerate,” which implies no moral weakness but instead refers to their origin. These genes are thought to be surviving relics of the ancient chromosomes from which the X and Y evolved, and they are identical to active genes on the X. The copies on the Y include 16 functional genes that are expressed throughout the body, where they perform housekeeping functions, while “many others are rotted-out hulks that no longer do any business,” Page has commented.

A third category of DNA sequences in the MSY region is “X-transposed.” These are bits of genetic code that jumped from the long arm of the X chromosome over to the Y, in a massive shift between 3 and 4 million years ago, after the ancestors of modern apes and humans split

Page says, “the Y chromosome has thrown overboard the vast majority of its genes, while today’s X has largely retained the 1,000 or so genes from the ancestral autosome.”

Is there a chance that the Y could be headed toward an ultimate total eclipse? According to Page, one can imagine the Y shrinking down to nothing but the sex-determining gene (which, after all, is how many scientists viewed it in the “rotting Y” models of the 1960s). Fortunately for the species, the Y has managed to preserve its package of spermatogenesis genes—some of which migrated from other chromosomes—that are cocooned among the amplicons. “The Y is doing a pretty good job of holding its own,” Page says.

LIGHTNING ROD

While the newly discovered gene-conversion strategy appears to be stemming the tide of gene decay, Page says research on that question is only beginning, and much remains to be done on many fronts. One of the top priorities is to sequence Y chromosomes from other men, and from males of other species, to garner new details about the chromosome’s evolution.

Mutations in spermatogenesis genes have already been implicated in forms of male infertility. Further analysis may lead to improved diagnosis and treatment.

Finally—and Page raises this point knowing it will be a scientific-political lightning rod—the finding of so much genetic information on the Y chromosome significantly widens the estimated genetic gap between men and women. “The difference between a male and a female comes down to trading the second X for a Y,” he says. “This trade involves about 1 to 2 percent of the genome, and that completely dwarfs all the genetic polymorphisms [normal variation between individuals] in the human autosomes”—all the chromosomes excepting the X and Y. Put another way, the genetic difference between a man and a woman is about the same as that between a man and a male chimp, or a woman and a female chimp.

This genetic reality will be difficult for many people to embrace and may only stir up greater controversy about genetic determinism, Page acknowledges. But he suggests that the disparities may underlie differences in disease susceptibility between the sexes, for example, and they should not be ignored in further research. **H**

