

## GENETICS



ISTOCKPHOTO

# Unraveling Our Own Code

## Nancy B. Spinner on the Latest Advances in Genetics

Interview by Andrew Plemmons Pratt

**ADVANCES IN GENETIC TECHNOLOGY** move very fast. The Human Genome Project, begun in 1990 and completed in 2003, finished two years ahead of schedule, offering the first complete look at the building blocks of human life. In just the past year, researchers have grabbed headlines by completing the first sequence that encompasses the genetic data from both parents of a single individual and creating an artificial bacteria genome from scratch. Numerous start-ups are currently gunning to capitalize on the promise of personal genetic sequencing, and a \$10 million prize from the X Prize Foundation has companies racing to be the first to sequence 100 human genomes in 10 days at a cost of less than \$10,000 per genome.

But apart from this rush to capitalize on the ability to sequence genomes, scientists are learning—at breakneck pace—about the connections between

the makeup of our DNA and the diseases that afflict us. A recent Associated Press article on the speed and scope of advances in the past few years quoted two Harvard scientists who wrote in the *New England Journal of Medicine*: “There have been few, if any, similar bursts of discovery in the history of medical research.” Understanding the directions and revelations of this work will be crucial to making informed policy decisions about the future of genetic medicine and genetic privacy.

New laboratory techniques allow researchers to peer into and control smaller and smaller sequences of genes. Nancy B. Spinner, a professor in the University of Pennsylvania School of Medicine Departments of Pediatrics and Genetics, focuses her work on identifying genes that contribute to disorders like Down syndrome, in part by considering the deletions and duplications across the genomes of vari-

ous individuals. Recent advances have revealed that genetic characteristics and susceptibilities to disease are not simply the product of individual genes. Rather, individuals carry different numbers of the same genes, and those extra copies—and in some cases, missing copies—of tiny genetic sequences can have serious implications for personal health.

Medical researchers may eventually be able to personalize the prevention, diagnosis, and treatment of diseases according to an individual's personal genetic makeup. Already, studies link gene variations with elevated risks of Alzheimer's disease and breast cancer. A study out just last week drew connections between genetic variation and schizophrenia.

Spinner spoke with *Science Progress* about this rapidly changing area of research and the challenges researchers face on the way to personalized medicine. This interview has been edited and condensed.

**ANDREW PLEMMONS PRATT, SCIENCE PROGRESS:**

The latest sequencing tools for genetics research are revealing huge stores of information that go beyond identifying just the functions of individual genes. Instead, there are significant patterns in the copies of genes distributed across a person's DNA and across a population. This is called "copy number variation." Can you explain this kind of research and how it changes the way scientists are approaching genetics?

**NANCY B. SPINNER:** First let me describe my own work, which straddles the borderland between diagnosis and research, because I run both a clinical diagnostic laboratory and a research laboratory, some of what I have to say may apply to one or the other. There have been a lot of huge advances in our field—it's been a seminal time for us. The human genome sequence is known; the first draft was deposited in 2001. And we have known since around that time that there are differences from person to person. This is not unexpected, of course, because we all vary. We are all identifiable and our sequences reflect that, although it's a relatively small portion of our genomes that vary by single base changes.

But starting in 2004, we've realized that we all vary not only by those single nucleotide changes, but that there are also larger regions that can vary in copy number between "normal" individuals, who do not have an apparent genetic disorder. This was really mind-boggling, because we've known for many years—since the late 1950s and early 1960s—that there are some people with clinical abnormalities or diseases, for example, Down syndrome, where there is an extra copy of a chromosome or a part of a chromosome. Since that time, there have been many examples where missing something or having something extra causes a disease that you can understand and study. But now, we have the tools to look at small pieces of chromosomes. When we look at whole chromosomes or parts of chromosomes, we're looking at millions and millions of base pairs. Conventional cytogenetic technologies require that at least 5 to 10 million base pairs of DNA be missing or extra before it can be identified. But the new technology allows us to identify deletions and duplications as small as 5,000 bases.

We are using this new information in two ways. My lab is really excited because we can identify smaller deletions or duplications that cause disease, and so far it seems that the new technology is doubling our chance of identifying a cause for the clinical problems seen in the patients who are sent to us for testing. But at the same time, we're finding small deletions and duplications in normal people. In some cases we can be very sure which abnormalities are "disease causing" and which are normal, but in many cases, we cannot be certain. This is a huge challenge for us right now. We are also trying to understand the function and implications of these "normal" variations, and how they impact the observable differences between us, both in terms of physical features, behavioral features and susceptibility to disease. This is a huge area of research right now.

**SP:** Besides Down syndrome, are there other diseases that your lab is focusing on in your research?

**SPINNER:** In our clinical lab, we diagnose all kinds of abnormalities that are caused by whole chromosomes being extra, as in Down syndrome. And there's a whole range of syndromes caused by deletions of parts of the genomes. My lab has focused on a disease called Alagille syndrome, which is a dominant disease that can sometimes be caused by a deletion of part of chromosome 20, and we've been studying this disease actually for 15 years now. We were able to identify the gene that causes Alagille syndrome, and we've been studying that gene. This is one of a number of different disorders caused by loss of something within the genome.

However, even in the case of Alagille syndrome, for which the disease gene is known, we're turning to studies of copy number variation to try to find out why some people who have this disease have a very mild form, while others have a very severe form. We're testing the hypothesis that some of the copy number variation that is found throughout the population might influence the severity of this disease.

**SP:** Let's go back to your remarks on the difference between "normal" and "abnormal" genes. Research over the last few years reveals a large amount of junk DNA in our chromosomes. How much information in our DNA may not be relevant to who we are?

**SPINNER:** Sure, though I won't say that anything is not relevant. This is one of the questions that geneticists are addressing. We know there are 3 billion base pairs of DNA. That's a lot of DNA that codes for about 30,000 genes. Overall, maybe 5 percent of that DNA actually codes for the genes. When I say "genes," I'm usually referring to a piece of DNA that codes for a product, usually a protein product. So there's much of the genome that we're not completely sure what it does. But now that we know pretty much the whole sequence, we can look at variation—both in the part that is known to code for genes and the stuff that isn't known to code for genes—and we're starting to learn what's in all

that stuff that we just didn't recognize before. So there is an ongoing project of the National Human Genome Research Institute called the ENCODE Project—or "Encyclopedia of DNA Elements"—and this is a consortium setting out to find out all the parts of the genome that are really crucial, that we just didn't recognize. We've gotten good at recognizing a gene. We know what its architecture looks like, and we can pretty much identify it in a rather straightforward fashion, but we don't know about all of the other parts of the genome and we're just starting to learn about those.

**SP:** There's been a lot of coverage in the mainstream press about people being subject to genetic discrimination based on predispositions they might have. Does this new information that is coming out of research on copy number variations have implications for policy discussions?

**SPINNER:** I think it has implications, and I think there is so much to learn that it's mind-boggling. Luckily, the threats of genetic discrimination haven't really been borne out in too many cases. The genetics community is happy to see that that hasn't been as much of a problem as I think people were worried about in the beginning. The problems that are now facing us are just at the very beginning of a) understanding the variation in the genome, and b) understanding how it relates to health and disease. It is all so new that it is very difficult. The science is just not there yet, to be able to tell us what it means to have something that is extra or missing.

I'll give you an example. We are now launching a clinical diagnostic test that looks at these very small duplications and deletions. There are many labs that are already doing it, and my lab is planning to start opening the doors to take in these tests on May 1. Many researchers have been applying this to many different populations.

There was a recent study in the *New England Journal of Medicine* on duplications and deletions on a small part of chromosome 16 that may be asso-

ciated with autism. This is one study. I don't think the data is convincing, but it is a published study in the *New England Journal of Medicine*. As we start running these tests clinically, we recently studied a patient—it was a research sample—and the sample was being done because the child had died of multiple congenital anomalies before birth.

So this was a serious, severe kind of abnormality. We looked at the parent's DNA, and in fact, the mother carries this duplication. There is a recent paper saying these same duplications can be associated with autism, but yet I had a family where the child had a totally different problem, had this same duplication, and the mother carries it.

So from my perspective, I'm very nervous. I don't want to overstate this relationship, I'm not convinced of it, but it's one report in a prominent journal. Anyone who Googles this will find this paper. This is just one example and shows you what happens. This is a brand-new tool. This is so new that the scientific community hasn't really had a chance to digest it yet.

We're very excited about these new tools for diagnostics, because in many cases, we're finding things we never could have seen before. The number of abnormalities that I'm convinced of has more than doubled with this new technology, so the number of patients for whom we'll be able to find diagnoses and explain what's happening and give them accurate recurrence risks is wonderful, but yet we don't completely understand the information.

**SP:** How important is government support of basic research in genetics? What are the possibilities for creating treatment breakthroughs?

**SPINNER:** Well, you're asking a scientist, of course, so I'm biased. Most of our work is funded by the National Institutes of Health. One of the reasons why I love what I do is that I sit at this boundary. I use basic science information all the time; I do diagnostic testing; a lot of my research tends to be translational in nature. So I'm really sitting on the fence between the clinical end and the basic science end. You absolutely have to have all of those pieces in place. There's absolutely no question about it. The basic scientists who work with model organisms like yeast or *Drosophila* or the mouse—they can really do experiments to try to understand. Once someone like myself identifies a gene that's important in a human disease, or a region of the genome, the only way to really study it is in a model organism. So you have to have all of those tools set up and in place to do that kind of very basic science work. I think that in the human genetics world, there have been some recent advances in a couple of different diseases, for example, Marfan syndrome and Fragile X syndrome, where the information we're getting is starting to be used for treatment. I think it's a really exciting time with a lot of possibilities. As a scientist, I think it's crucial that we fund this work. We have to fund the people doing the health-related research, so they can identify the problem, and you have to fund the basic scientists. **sp**