AUTHOR COMMENTARIES - From Special Topics

Tuberculosis - January 2009
Interview Date: March 2009

PDF

Stewart Cole
From the Special Topic of Tuberculosis

In our recent Special Topics analysis on Tuberculosis (TB) research over the past decade, Professor Stewart Cole ranks at #2 by total citations, based on 51 papers cited a total of 4,662 times. Professor Cole is also the lead author on the most-cited paper overall in the analysis. His record in Essential Science Indicators℠ *from Thomson Reuters includes 116 papers cited a total of 6,764 times between January 1, 1998 and October 31, 2008. He is also a Highly Cited Researcher in Microbiology.*

Professor Cole is the Director of the Global Health Institute as well as the Chair of Microbial Pathogenesis in the School of Life Sciences at the École Polytechnique Fédérale de Lausanne (EPFL) in Switzerland.

Below, ScienceWatch.com *correspondent Gary Taubes talks with Professor Cole about his TB research.*

SW: What prompted your initial research interest in TB?

I was working on leprosy originally, and there are quite a lot of similarities between the leprosy bacillus and the tuberculosis bacillus. Even in the mid-1980s, there were quite a lot of cases of multi-drug resistant TB, which was becoming a major problem in the industrialized world again. So I decided to make a sideways move and get involved in TB research.

SW: As first author of the 1998 TB sequencing paper in *Nature*, your most-cited article, it suggests you were the driving force behind the project. Is that the case?

Yes. I knew we had the material in hand to sequence the TB genome. We had all the biology and the genomic DNA and ordered libraries and clones, which could be used for sequencing or mapping. So I went around to different funding agencies to try and convince them to support the project and put up the necessary cash. Some of them said this kind of stuff is not important, or that this wasn't real science, but in the end I managed to convince the Wellcome Trust to fund the bulk of the work.

SW: Considering your difficulty getting funding, was it challenging, as well, recruiting a team to analyze the data? And how did you decide whom to approach?

It wasn't difficult at all to convince the scientists to get involved. Most of them could see immediately that this was hot stuff, something that would turn out to be important. The funding agencies were just much more conservative. As for whom we contacted, we wanted people with expertise in different areas. Obviously, we wanted people, for instance, who were strong in genomics. That's where the Sanger Institute came in. They played a really great part.

To understand the biology, we had to recruit experts, specialists, in very unique areas of science—lipid metabolism, for instance—and so we asked Clif Barry to get involved and give us a hand and analyze the data. He played a fantastic role in the project. We also recruited experts in membrane proteins and transcription factors. You have to realize that these genome projects generate colossal amounts of information, far more than one mind can possibly handle or analyze, so teamwork was critical. Certainly getting the right people together was part of the challenge, but it was also very rewarding.

"The vaccine work has progressed quite a lot thanks to the genome."

SW: Were there other significant challenges that had to be overcome in doing this research and reporting the results?

What made it quite an effort to write that paper was that we had so much to say.
These days, genome papers are fairly trivial, because there are such a lot of them. I think our paper was maybe the 11th or 12th genome published; it was the first one for a major pathogen and we had a lot of new information. The difficult thing was to decide what was important and worthy of inclusion in the paper, and what was peripheral. At that time it wasn't really obvious.

So we focused on a number of areas that were highlighted by the abundance of the genes in the genome. We found a huge number of genes involved in lipid metabolism, for instance, both in the synthesis and the degradation of lipids. So quite a lot of the paper is devoted to that.

Another thing we found was a whole series of genes or gene families coding for unusual proteins that had never been seen anywhere else. They turned out to be surface proteins of *Mycobacterium tuberculosis* (MTB). That was quite an important finding, and so we devoted considerable space to it.

Then, as the title says, we learned an awful lot about the biology of the organism from its genome. The sequence enabled us to predict which metabolic pathways were present and make some informed guesses about the physiology of MTB, and quite a lot of those guesses turned out to be correct.

SW: Were you surprised at how much you learned?

No. I don't think so. I was always a believer in genomics as a way of generating large amounts of new information. I was surprised, though, by some of the things we found. It turned out that there were a lot of protein kinases, for instance, that were more similar to eukaryotic kinases than prokaryotic ones. That was quite an interesting finding. There are a lot of surprising vignettes like that.

SW: If you published your paper again, what, if anything, would you write differently and why?

I think we got the balance right. I think we made the right decisions at the time. A lot of things that we said were interesting, particularly in the lipid metabolism area. We predicted how some of the glycolipids of the cell envelope would be made, and those predictions turned out to be correct. We also predicted how some of the lipids would look; we speculated that they might be important, and subsequently other investigators confirmed that.

In doing this, I think the paper played two major roles. One was in making available a lot of this new information to the community from a very early stage, and the other was in speculating and presenting some new hypotheses, which a lot of other people went on and tested. We generated such a huge amount of information from the genome; obviously one laboratory alone can't possibly handle all that.

SW: Your second most-cited paper is the 2002 *PNAS* article, "A new evolutionary scenario for the *Mycobacterium tuberculosis* complex." What was the new evolutionary scenario and why has this paper been so influential?

It's always been claimed that TB was a disease of zoonotic origin and that humans had contracted it from infection with a cattle form of the bacillus -- from *Mycobacterium bovis.* What we reported in that paper, from work derived from the genome sequence, is that we identified a series of polymorphic markers, in particular regions, which had been deleted from the genome of some species—such as *Mycobacterium bovis* or BCG (Bacille Calmette-Guérin), the strain used in vaccines—and found that these deletions had occurred after *Mycobacterium bovis* and MTB had separated.

So, in fact, MTB was the ancestral strain and *Mycobacterium bovis* was descended from it. This suggested that the theory in the literature was wrong. Rather than humans acquiring the disease from cattle, it looks more likely that cattle acquired the disease from humans.

SW: Has this new evolutionary scenario held up with further research?

Yes, that's been confirmed by other investigators and with different techniques. This turned out to be a remarkably robust and accurate model.

"I knew we had the material in hand to sequence the TB genome."

SW: How has the TB research landscape changed in the decade since you sequenced the genome?

I think people have been able to do experiments in a much more informed manner. We can see what genes are in the genome, and now in about half the cases we know with great accuracy what their function is. People are now able to do informed experiments: they have a hypothesis and they can test it in a direct way, rather than indirect ways, as in the past.

It used to be that people would try to assess the effect of a gene by isolating mutants. When you can see what the gene does, you can test it in a forward manner. It's been amazingly helpful and extremely useful for biochemists, enzymologists, and crystallographers, those who work with proteins. They can take genes, express them in bacteria or yeast, purify the recombinant protein, test its activity, and so on.

SW: Did the genome sequence help in developing a better TB vaccine?

The vaccine work has progressed quite a lot thanks to the genome. Again, for instance, to illustrate my point, people have always been aware that important proteins for generating an immune response were often secreted proteins or surface-exposed proteins. With the genome information in hand, we could identify those proteins just from looking at the sequence. Then we could draw up a hit list of interesting proteins to test and express. It's a question of doing things in an informed manner rather than blindly.

SW: What is your research focusing on today?

Mostly drug discovery for TB. In this field the genome information has been extremely useful also. First of all, as I said, it provides us with information about potential drug targets, so we can test hypotheses. Secondly, thanks to some of the new technologies, like gene expression microrarrays and transcriptomics, we can obtain information about how existing drugs or even new compounds work by studying which genes are turned on and off in response to treatment with that drug. This has been quite a useful tool.

Thirdly, the genome has been really useful for identifying targets for new drugs by means of isolating resistant mutants and then sequencing the genome to find where the mutation is. This has been quite a fantastic gain in time in terms of identifying drug targets. Traditionally people would have done that using genetic approaches, but MTB grows really slowly. It used to take six to nine months to find that information, whereas now we can get that information in a week.

SW: What do you consider the most challenging aspect of TB research?

It's not impossibly difficult to work with MTB, but two factors are major problems. First, as I said, it grows very slowly, so experiments take a long time and that can be rather frustrating. An experiment you can do with *E. coli* in a week, might take six months to do with MTB. And then, of course, the other major constraint is that MTB is a category 3 pathogen. So we have to work in containment facilities, using very stringent safety procedures. That's a major restriction and not every research center has these facilities, so that really slows things down.

SW: What unexpected or serendipitous events arose in the course of your research?

Nothing that I would call unexpected or serendipitous, but where we got lucky, if you like, with the genome is by getting the right people together, putting together a team that worked well and efficiently and with whom it was really enjoyable experience to work. Often times when you work with people you don't know, it can turn out to be less than a great experience. Doing the genome project was an extremely positive and memorable experience. A lot of the people we worked with, some of the authors on that *Nature* paper, I had never met until then. They played a really important part in the project and it was rewarding to work with them.

SW: If we lived in an ideal world, and you had an unlimited source of funds to do one single experiment, what would you do?

Well, this is a little bit disconnected from the topic of our discussion, and it's not exactly an experiment, but it would be to invest in technology that would show whether a candidate drug would actually work in humans. That would be an important step forward, because we can't really understand what compounds are going to cure diseases in humans and which ones won't. We're starting to build a basis of understanding, but it's still pretty fragmentary. If we had a good predictor of what will work and what won't, that would very much help drive the science of drug discovery.

Professor Stewart T. Cole, FRS

Global Health Institute
School of Life Sciences
École Polytechnique Fédérale de Lausanne
Lausanne, Switzerland

Stewart Cole's current most-cited paper in *Essential Science Indicators*, with 2,752 cites:

Cole ST, *et al.*, "Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence," *Nature* 393(6685): 537-+, 11 June 1998. Source: *Essential Science Indicators* from Thomson Reuters.

Additional Information:

Stewart Cole is featured in ISIHighlyCited.com

KEYWORDS: TUBERCULOSIS, TB, MYCOBACTERIUM TUBERCULOSIS, MYCOBACTERIUM BOVIS, COMPLETE GENOME SEQUENCE, LIPID METABOLISM, SURFACE PROTEINS, METABOLIC PATHWAYS, GENOMICS, BACILLE CALMETTE-GUERIN, EVOLUTIONARY SCENARIO, VACCINE, DRUG DISCOVERY.

PDF

back to top

Special Topics : Tuberculosis : Stewart Cole Interview - Special Topic of Tuberculosis

Scientific Home | About Scientific | Site Search | Site Map