A conversation with Francisco Mojica, Alicante, Spain

"I'm a Microbiologist, I Love Basic Research"

Who invented CRISPR is a ferociously-debated question. However, we asked: who *discovered* CRISPR? That was Francisco Mojica from the University of Alicante in Spain – and he also coined the term CRISPR.

In 1987, Yoshizumi Ishino of the Osaka University was the first to observe repeated sequences with defined spacing. But you are named as the discoverer of CRISPR. How can that be explained?

I started my PhD in 1989 looking at the archaeon *Haloferax mediterranei* that lives in salty ponds nearby in Santa Pola. We did some sequencing. In fact, these sequencing experiments were among the first ever performed at my university back then. I discovered very peculiar repeats: they were 14 perfectly identical repeats, each 30 base-pairs long, regularly interspaced by variable sequences. They were palindromic repeats and they were transcribed. Of course, there were repeated sequences all over the genomes known at that time but such repetitive structures had not been described, except by Ishino. However, *E. coli* is an enterobacterium, phylogenetically far distant from *Haloferax*. And we later also found similar repeats in *Haloferax volcanii* and in the genus *Haloarcula*. So we thought: these sequences might be relevant and play a fundamental role in prokaryotes.

"The reviewers of our grant applications told us that we need to work with an established model organism."

But you had no idea what role that was?

(laughs out loud): No, absolutely not. First, we thought the sequences might be involved in recombination or replicon partitioning. But we could not find any clear evidence. In 1995, I went to Oxford for a postdoc and worked on the influence of DNA topology on gene expression in *E. coli*. I did a lot of experiments with a protein named H-NS. Aside from that I always worried about the repeats, looked for related literature and tried some experiments to find out whether these repeats have anything to do with DNA topology. Obviously, they don't. When I went back to Spain and got a position in Alicante, I started my own research group. We tried many hypotheses but these structures kept being enigmatic. We put *Haloferax* aside and went on working with *E. coli*.

"We were not able to convince people that we had made a great discovery."

Why was that?

Because the reviewers of our grant applications told us that we need to work with an established model organism – and *Haloferax* wasn't any kind of that. Looking back, that was a double-edged demand that took us a lot of time but also directed us to the right scent.

What do you mean by that?

Well, eventually, we amplified and sequenced the loci of *E. coli* and found that one of the individual spacers between the repeats was identical to some sequence of a phage that could infect *E. coli*, but not that special strain we were working with. That was totally resistant. So we thought: wow, what's that? In the 1990s, more and more sequencing data were published. We looked at them in detail, as far as they were available, with a programme that one of our group members had developed. We came up with many false positives but we also identified repeats in other prokaryotes and even a sequence in a plasmid in mitochondria of *Vicia faba* – the common bean. We published the work in 2000 and argued that based on their wide distribution, these repeats should be of ancestral origin and should have some sort of biological role. This paper was followed two years later by another description of repeats, identified by Ruud Jansen of the University of Utrecht. We both had chosen different names and agreed then on calling them CRISPR, clustered regularly interspaced short palindromic repeats.

Didn't actually you *coin that term?*

(laughs): That's right.

Ángeles Gómez (Feedbackciencia.com)

Coming back to E. coli *– how did it hamper your work, as you mentioned earlier?*

We identified spacers in many of the strains that were sequenced at that time. Two percent of them were almost identical to sequences of loci in plasmids and viruses. We then realised that it was never reported that strains with certain spacers contained harbouring plasmids or viruses with the corresponding sequences. So we thought: these spacers could probably be derived from sequences of genetic elements that invaded the ancestor of that strain. Could that be some sort of an immune system? We tried to prove that hypothesis experimentally with *E. coli* but we did not get reliable results: sometimes it worked, sometimes it did not. That was really frustrating. Today, we know the reason: in most *E. coli* strains we were working with, CRISPR is repressed. And you know what? Of all molecules, it was the histone-like protein H-NS being responsible for that. Isn't that full of irony?

Did you get your ideas published?

Not so easily. While the experiments with *E. coli* weren't working, we went back to the literature on the archaea and found the same resistance pattern for methanogenic and sulphur-reducing archaea. Infectious elements were not able to get into spacer carrier strains. We tried four journals to publish these findings, but they all rejected our paper. That was really very much disappointing. We were obviously not able to convince people that we had made a great discovery. The reviewers asked for more experimental proof but because the bacteria didn't work, we had none. Eventually, we got it published in *Journal of Molecular Evolution* in 2005.

When was your hypothesis finally proven?

That was in 2007, by researches from the Danish company Danisco and the group of Sylvain Moineau in Canada. They worked with *Streptococcus thermophilus* that constitutively expresses CRISPR. They found that this bacterium becomes resistant to phage infections when they introduced phage-matching sequences as spacers in the CRISPR locus. That was the ultimate proof. On the one hand, that finding was very frustrating again because we would have liked to find such evidence ourselves. But on the other hand, that paper was great because finally, our hypothesis had to be accepted as being really correct. They also proved that the Cas proteins, that Jansen had identified earlier, were a functional part of the CRISPR system. So the CRISPR-RNA guides the cleaving Cas proteins to their homologue in invading sequences, similar to the mechanisms of interference. That's it.

Bacteria also have an innate immune system. In an evolutionary context, which system is older: innate immunity or CRISPR?

Arguably, innate immunity. But CRISPR is, for sure, very ancient; most likely 3,000 to 3,500 million years old. Remember what I said about spacers in the mitochondria of *Vicia faba*? These are almost identical to the canonical sequences of cyanobacteria. So, to me it is clear that they come from the ancestor of cyanobacteria. Perhaps the CRISPR system is older than the first eukaryotic cell.

"Perhaps the CRISPR system is older than the first eukaryotic cell."

Today CRISPR-Cas has become so public because it serves as a very useful genetic tool. Have you been working on engineering that tool?

No, no, no. Sure I'm happy that it is so useful now – that so many tools have been developed. And I'm proud of being part of the very beginning of all that. But I'm a microbiologist, I love basic research, love understanding how biological systems work and I'm proud of that also.

So you're not by any means involved in that patent dispute?

No, not at all.

What are you working on now?

I kept working on CRISPR. We characterised a few elements, we described and defined the PAMs, the protospacer adjacent motives, that had been initially found in 2005 by another group. We realised that the PAMs are common features in most CRISPR systems. And we looked for factors that influence the CRISPR activity.

"This story tells us, exemplary, how important it is to support fundamental research."

Have you banned E. coli *from your lab?*

(laughs): No. You know that we found an anti-Cas system?

No, tell us.

We saw that the majority of *E. coli* has an anti-Cas system. They carry only very small CRISPR arrays and the spacers therein are quite similar to sequences of *Cas* genes that are actually missing in the bacteria. When we found that, a few years ago, we thought that the CRISPR remnant is a system to prevent the acquisition of new Cas. Now, we have demonstrated that we were right. We introduced *Cas* genes into such *E. coli* strains. The Cas encoded by the plasmids were loaded with CRISPR-RNA of the remnant array and directed against these Cas-containing plasmids.

Isn't that a case of a bacterial autoimmunity?

You could say so. The bacteria lost their *Cas* genes and by the remnant CRISPR, they make sure that they don't get them back.

And what's the biological rational for not having a functioning CRISPR system?

CRISPR is probably the first line of defense in *S. thermophilus*. If you challenge it with a virus, a few survivors can be found and the vast majority of them have acquired new spacers from the phage, and therefore acquired immunity to that phage. But only fifty percent of bacteria have the system and, in most cases, they have it repressed. We found that in *E. coli*, the complexity and function of CRISPR negatively correlates with their pathogenicity. That means: the more pathogenic, the less number of spacers and repeats. Also, most pathogenic *E. coli* strains don't have any CRISPR-Cas at all. Meaning: CRISPR is not good for them. How can that be? There must be some drawback. We think the prize is that a cell having CRISPR cannot gain foreign DNA. The system prevents horizontal gene transfer. However, that might be important for bacteria for surviving.

The main example is the acquisition of antibiotic resistance. These strains are quite pathogenic. With their anti-Cas, they make sure that they can introduce new DNA, such as new virulence factors or other genes, necessary for high virulence. They take the risk to be infected by viruses, instead of being immune against invaders. Perhaps, they are not challenged very frequently, so fighting against viruses and plasmids is not a big problem for them.

Would you like to have a last word?

This story tells us, exemplary, how important it is to support fundamental research. You never know what you will find and what it might be valuable for. That point should never be forgotten.

Interview: Karin Hollricher