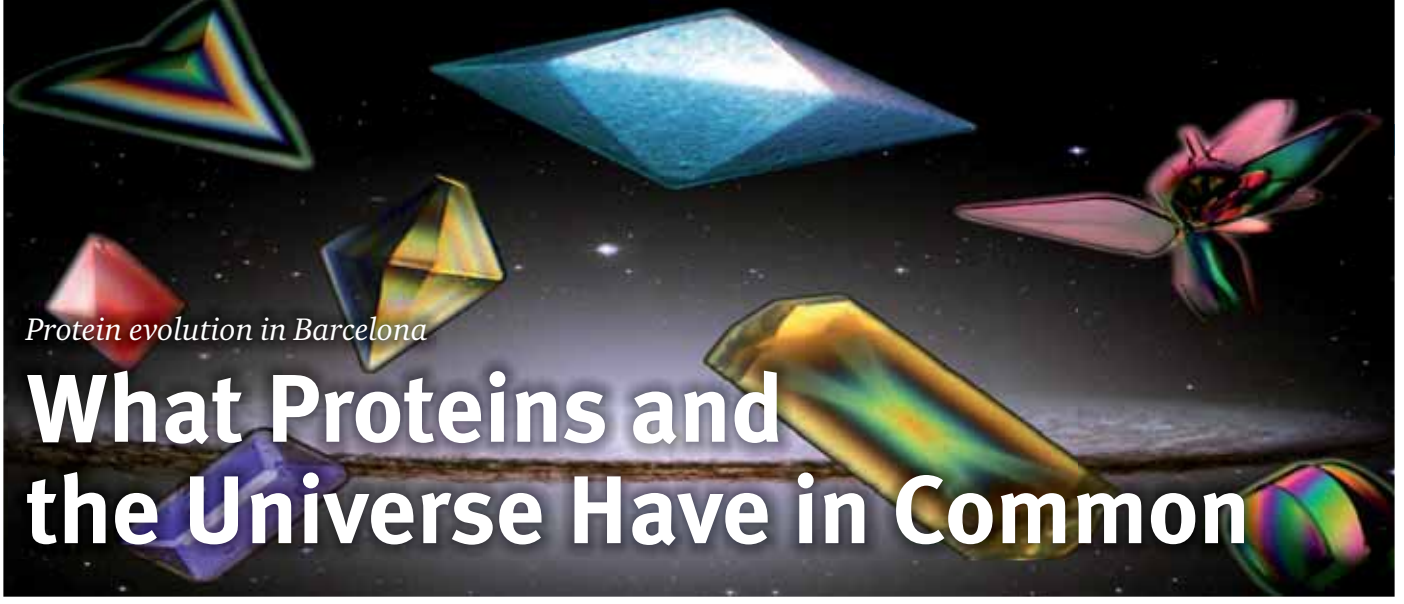

Protein evolution in Barcelona

What Proteins and the Universe Have in Common

Proteins are still diverging from their consensus sequences, as two Russian-born computational biologists have just observed. This way, an expanding sequence space is formed, similar to the ongoing expansion of the universe.

The idea that we live in an expanding universe is one of the most unexpected and important discoveries of 20th century science. It was Edwin Hubble who, in 1929, put forth this suggestion as the expanding universe theory. Almost 100 years later, Inna Povolotskaya and Fyodor Kondrashov at the Centre for Genomic Regulation (CRG) in Barcelona, Spain have applied his approach in answering key questions regarding the evolution of proteins. Their corresponding study entitled 'Sequence space and the ongoing expansion of the protein universe' has just been published in *Nature* in May 2010 (*Epub ahead of print*).

Lab Times spoke to both authors, Inna and Fyodor, and not only learned about fitness landscapes and sequence spaces, but also got some 'behind the scene' insights into how this study was carried out. Fyodor heads the Evolutionary Genomics group at the CRG and Inna is a research assistant soon to embark on her doctoral study within the same group.

Fyodor explains how the idea for this study first took shape, "One of the main questions that interests us is the structure or shape of fitness landscapes in nature. We started out by asking whether or not proteins with conservative functions are located on narrow fitness peaks. First we set out to describe the distribution of proteins in sequence space. Then, following a hint from Danny Tawfik from the Weizmann Institute in Israel, we tested whether or not sequences diverge from their consensus. The idea was that a sequence consensus is broadly relevant to the centre of the fitness peak."

A journey through protein space

Inna and Fyodor compared protein sequences from different species and measured the evolutionary distance between them. In addition, they devised a computational method for calculating the pace at which proteins accumulate amino acid substitutions. One has to keep in mind that these substitutions are often restricted in order to maintain the structural and functional integrity of the protein. Just as Hubble correlated the distance between galaxies with the velocity at which they moved away from each other, the CRG researchers could show a similar association by comparing the evolutionary distance between proteins and the rate of their divergence.

Large amounts of data

One can imagine the large amounts of data generated by such an analysis. "The programming part was not as complicated as the scripts were quite straightforward," says Inna, who has a background in physics. "While analysing such a large volume of information, we would often end up with a few terabytes of data, and to handle all that was quite the challenge." However, the young researcher found the entire experience quite fulfilling. She explains, "It was fun because the entire project was like a quest from beginning to end. We would generate the data, interpret it and this would decide which step was to be taken next."

Fyodor Kondrashov (l.), Inna Povolotskaya

The truth about protein evolution

Once the whole analysis was complete, some very interesting results surfaced. The scientists were able to conclude that even after an evolutionary period of 3.5 billion years, ancient proteins still continue to diverge from each other. Simply put, the limit of protein evolution is yet to be reached. They also observed that this rate of divergence was slow and that "the limit of sequence similarity imposed by common function may not exceed that of random sequences".

Fyodor explains how their observations correlate to Hubble's approach. "Once we saw that extant proteins are still diverging from their consensus, we realised that this observation is broadly equivalent to the expansion of the real universe and that we could observe this expansion without looking at a hypothetical 'centre' of the sequence space, but we can, just as Hubble did with galaxies, rather look at the divergence of extant proteins from each other."

After the 'Big Bang'

Okay, let's do it that way. Whereas physical space continues to expand after the Big Bang, sequence space is a pre-defined abstract entity, a limited subspace, which corresponds to sequences with a specific function. In both cases, the correlation between distance from one point of observation to reference points in space and the relative rates of divergence reveals whether or not these objects (proteins or galaxies) recede from a common point of origin or have reached the limit of their divergence.

"I think that perhaps the most striking result of this study is that even the most conservative protein functions can diverge

in sequence to a very low level of similarity," he adds. And what do the researchers plan next in this direction? "There are several things on which we can expand. Perhaps our approach can be applied to proteins or domains that have either evolved very recently or have diverged beyond recognition from other sequences, in the hope of resolving this issue," elaborates Fyodor.

Bringing together bench and workstation

"Our group at the CRG is interested in practically all aspects of evolutionary biology. I really like to work on the interface of several disciplines and to combine both holistic and reductionist approaches. On a day-to-day basis we tend to work bioinformatically, although currently we are expanding into our own simple experiments. We may be one of the few laboratories where wet-lab work is used as a means of support to our bioinformatic queries and not the other way around," says Fyodor.

Both researchers are very satisfied with the research environment at the CRG. Inna feels that the facilities at the institute are of high quality and this fuels independent innovation. Fyodor elaborates, "The CRG is a very strong institute, second to none. What ticks for me is a combination of hard work and lack of stress. Everyone works very hard here but in between the work people are relaxed. Any student can expect to be challenged here and work with a young and international group of motivated scientists."

Science as a passion vs. a profession

On being asked why she chose to pursue scientific research, Inna replies. "I like how one can work on what one finds interesting, and address relevant questions. I enjoy how my work on a certain topic can connect to many different questions. Today, I could be working on one topic and one year down the line I could be involved in something else."

For Fyodor, this decision was a socially-determined chance. He explains, "I spent a lot of time in my father's lab back in Cornell, when I was still in high school. In college I considered social work as a career but a random chance put me in the lab of Eugene Koonin and I did not look back. For me, science is like entertainment. I really do not think of it as a job. I think about questions and answers that entertain me and I work on those. Of course, this means that I tend to skip between questions but at least I feel like my lab being a huge and complicated playground." And Fyodor closes, "In principle, I would like to integrate some real field biology with genomics. I feel that genomics is becoming cheap enough to take molecular evolution from the lab back into the field."

Inna undertook her formal education in the faculty of physics at the Moscow State University. She graduated in 2008 while specialising in medical physics. Bioinformatics was a field that had always provoked her interest and this prompted her to join Fyodor's group in Barcelona. Says Inna, "I had heard of Fyodor and his work and I knew that I could learn from him."

Informal education

Fyodor's scientific education was deeply influenced by his family background and upbringing. He elaborates, "I come from a family of scientists and my parents, grandparents and siblings are engaged in some intellectual query of the natural world. Most of my education I attribute to informal work, either with my father or with Eugene Koonin. Formal education in school and graduate school was useful and interesting but contributed only a small fraction to my development as a scientist.

And how do the two scientists relax at the end of a hard day's work? Inna enjoys travelling, and visiting new places. "I like how convenient it is to travel in Europe. On weekends and long holidays, I am almost never in Barcelona," she adds with a smile. Fyodor's side of the story is quite different, "Since I got a lab, a bit over a year ago, I don't recall having any free time. A switch from the lab into fieldwork sometimes feels like free time," he adds.

Informatics, an elegant tool in research

Inna's and Fyodor's current study clearly reflects how informatics can be applied to answering very fundamental questions about the living world around us. The key step in this field is not just framing the right questions but also being able to devise the best path to get to the answers. Both, Inna Povolotskaya and Fyodor Kondrashov, are shining examples of how this can be achieved.

In conclusion, a famous quote comes to mind "Equipped with his five senses, man explores the universe around him and calls the adventure Science." These words were uttered by none other than Edwin Hubble himself. In today's world, informatics could very well be thought of as a sixth sense that is guiding our attempts at better understanding our universe. LATIKA BHONSLE