



# Extreme conditions leave universal mark on extremophiles

## Singing from the same songsheet

19 January 2024/ Interview with Lila Kari, University of Waterloo

Part of the show [The origin of Europe's MS disease, and South Pole sequencing](#)



*Credit: CC0*

There are plenty of extremes out there: high and low temperatures, high acidity, high pressure, and surviving in them also requires some seriously specialised genetic adaptations. And it turns out that the wide range of these extremophiles, microorganisms

**capable of making a living in hot springs or deep sea trenches, might have more in common with each other than previously realised. The University of Waterloo's Lila Kari explains...**

Lila - We are looking at extremophiles, which are organisms that not only habit but thrive in environments characterised by extremes. They are found in the most inhospitable places on earth, including volcanoes, underneath polar sea ice, hydrothermal vents on the seafloor, and even in the presence of radiation and toxic waste.

Will - So to find out just what makes them so hardy, you've been having a look at their genomes. How exactly did you do that?

Lila - So we looked at the genome of extremophiles, organisms in general, in terms of a language to clarify. In the same way we use the letters of the Latin alphabet to write text and bits zero and one to write computer code, the four basic DNA units used by nature to write genetic information as DNA sequences. And what we studied is what is called the genomic signature of DNA, which is obtained by counting the occurrences of DNA words in a DNA sequence randomly selected from an organism. And here by a DNA word, I mean a sequence of DNA letters, for example, cat is a three letter DNA word consisting of the letter CAT. So it turns out that word frequencies are important and contain taxonomy information about the organism, species, family, genus class, and so on. And this is akin to being able to tell apart an English book from a French book by the fact that the English book is going to have a very high count of the word 'the', and the French book is gonna have a very high count of the word 'les'. So you can tell them apart by counting words and looking at the word usage profiles without knowing a single word of English and French. And in the same way we can study DNA in the genome by counting frequencies without needing to know anything about genes, proteins, or anything like that. So this method has been proved very successful for biodiversity species identification and species classification. Just look at the word counts.

Will - This seems like an extraordinarily large amount of data you had to sift through. Presumably you had help from cutting edge technology.

Lila - Actually, no, we don't need the entire genome. Remarkably you can take any short DNA fragment. The human genome is 3 billion letters. But we can look at a segment that is maybe 5,000 letters and by word counts of this short fragment, we can determine the species. And this genomic signature is so stable that it doesn't matter where you expect it from, from the beginning, the middle, the end, chromosome one, chromosome five, gene, non gene, long sequence, short sequence, they all have the same pattern. So this is one of the strengths of the method. You do not need to sequence the entire genome, you just need that tiny little fragment.

Will - So what did you do once you'd found this out?

Lila - So this has been used very successfully for taxonomic classification. But we looked to see if this genomic signature contains other information, for example, environmental information. And we knew that if there was any hope to find an environmental signal, we would have to look at extremes. And this is why we decided to look at the genomes at extremophiles.

Will - Does that mean then that if the environment in which something lives in is so extreme, it will leave a mark on the genetic information of the thing that's living in it?

Lila - Yes. That was our hypothesis, and it turned out that we confirmed that, as unlikely as it seems. And for that, we used the machine learning methods, both supervised machine learning and unsupervised machine learning.

Will - So does that mean then that if two organisms that may be unrelated live in similar conditions, that their genetics might be the same?

Lila - Their genomes certainly are the same. So what we found is that, as unlikely as it sounds, as unlikely as is, two organisms, a bacteria and archaea that are more distantly related than ebola is from a lichen, were grouped together as similar in terms of word usage because they are both adapted to high temperatures, for example. And this environmental signal, it's pervasive, it's everywhere, sprinkled everywhere along the genome, and you can detect it everywhere. No matter where you take a DNA fragment in the genome. It's a little bit like finding a new dimension of the genome. You know, we thought it was only text and actually it's songs and it has also a musical signal besides the lyrics and the text in it.

Will - So if you flip that on its head, then if you were to just look at a genome in abstract, you might be able to tell where it lives just by the frequency of certain areas of it.

Lila - Absolutely. Yes.

Will - That's quite extraordinary. So does that mean that we could potentially use this in a way of mapping organisms throughout the world?

Lila - Yes. However, I would say I am not sure that this applies to normal environmental conditions. Remember, we looked at really extremes like organisms that are very, very hardy and they have to live under very harsh conditions for a very long time. Probably for like more normal conditions, the signal is more faint and more work needs to be done in order to be able to see whether we would be able to detect. I'm sure the signal is there, but whether it is too faint to detect for normal environmental conditions, that remains to be seen.

Will - But that could well be where you'd hope to go next in this study then.

Lila - Yes, actually it's very interesting. We are also interested in this because people are interested in space missions and Mars missions and outer space. So some of these extremophiles, for example, *Deinococcus radiodurans*, with a radio tolerant organism, it survives vacuum radiation, desiccation, cold temperature, you name it. It was proved not by us, by other scientists, to be able to survive outer space for one to three years. So very interesting questions arise regarding what it takes to be able to survive out in space. And as it turns out, it's not enough to change this gene here or this protein here. What you need is like a pervasive change along the entire genome. So it's not as simple as you think it is.

Will - I would never assume it would be anything close to simple when it comes to sequencing genomes. So throughout this, were there any highlights or any organisms that you'd love to shout out as your favourite?

Lila - Well, I have to mention my favourite extremophile, which is *Pyrococcus furiosus*. You've got to love the name. *Pyrococcus* means ball of fire and *furiosus* means furiously. And it's called that because a little bowl of fire swims furiously in a hot aquatic event of 100 degrees temperature, which is its optimal growth temperature.

---

©The Naked Scientists® 2000–2020 | The Naked Scientists® and Naked Science® are registered trademarks created by [Dr Chris Smith](#). Information presented on this website is the opinion of the individual contributors and does not reflect the general views of the administrators, editors, moderators, sponsors, Cambridge University or the public at large.