Research Strategy: Christophe Bonny

Drug discovery in the age of deep sequencing

It took time for our microbiome, and in particular, our gut microbiome, to generate interest among scientists. Originally perceived as simply part of the body's waste system, it is now recognised as an active component of our metabolic health and an agent of our mental health through its immune and inflammatory responses to internal and external stimuli.

Today, the microbiome is best described as an independent organ of the body composed of an estimated 1.5 kilograms of bacteria. It is in close contact with, and therefore potentially capable of modulating some 70% of our immune cells, and many of the hormone-secreting bodies in our endocrine system. It also has the potential to influence behaviour through the gut-brain axis.

One of the most exciting aspects of the gut microbiome is its genetic diversity. Deep sequencing of the DNA in our stools – DNA that mostly originates from our intestinal bacteria – has revealed thousands of novel bacteria species. Our company, Enterome SA, has shown that these organisms code for some 21 million proteins, a figure that we obtained by aggregating the DNA sequences from the stools of 26,000 individuals and translating them into distinct proteins. By distinct, we mean proteins that are capable of carrying out significantly different functions. Knowing that the microbiome acts as an organ, impacting other organs, our goal is to identify novel drugs by characterising the secretory products, such as peptides and proteins, that come from this organ.

In this article, we discuss the specific approaches we have taken to mine this resource of 21 million proteins, assign functions to the relevant ones, and progress clinical drug candidates. This is done, in the first instance, by creating a database of microbiome-based proteins.

Estimating the diversity of the microbiome

How diverse is the repertoire of these proteins? One way to estimate the overall diversity of the human gut microbiome is to count the number of new proteins that we identify each time we deep sequence and translate DNA in the stools of a new individual. We have discovered that the number of new proteins identified reaches a plateau at approximately 20,000 individuals sequenced. This means that the relative contribution of every additional individual to the total number of proteins identified is close to zero. 26,000 individuals is well into this number so it is fair to assume that 21 million proteins captures a large proportion of the overall protein diversity of the microbiome. If a significant number of new proteins were to be identified, they would certainly come from the sequencing of new groups within distinct geographical locations. Individuals in the Enterome database are geographically spread from around the world. Interestingly, the vast majority of proteins that we have characterised are rare, which is to say they are not shared among most of us.

While it is accepted that the composition of our individual microbiomes is a key contributor to human health and has an impact on an individual's response to a treatment, 21 million is a huge number of proteins to investigate in order to isolate the critical ones. By comparison, our own human genome codes for 23,000 proteins and we have yet to fully understand the complex network of interactions among them. Also, if we look at the number of potential antigens presented by the proteins in our microbiomes, it becomes apparent that all possible human antigens are fully replicated within the 21 million microbiome-derived proteins. In other words, each time we look at a specific human antigen, for example an autoantigen governing an autoimmune disease, there is usually a group of individuals somewhere that has a perfect or close to perfect replica ("mimic") of the antigen expressed in their intestines and in close contact with their immune cells. In fact, different studies have shown that some patients with autoimmune disease also display intestinal "mimic" antigens.[1]

Large microbiome-based association studies have become the common approach for characterising the genetic elements which code for a given phenotype in an individual. The concept and indeed prerequisite here is that different individuals sharing the same phenotype – for example the same response to a treatment – would have the same or very similar microbiome bacterial species coding for the same or a similar set of proteins. While this is certainly true in a number of instances, it is clearly not always straightforward.

For example, we and others who have access to large cohorts of checkpoint inhibitor-treated cancer patients, have not been able to identify any species or consortium of bacterial species associated with treatment outcomes.[2] However, what we have learned is that while it is clear that the composition of the microbiome does have an impact on patients, the individual composition of the different microbiomes might deliver efficacy through different targets or pathways. This might explain why it has been difficult to assign one specific clinical response to one or a limited consortium of bacterial species or strains.

Looking at the rare protein

In many cases it might therefore be fruitful to investigate the large and thus far, mostly untapped repertoire of rare proteins produced by our microbiomes. These are the ones that escape characterisation by classical genetic linkage studies. To put this into perspective, and assuming a reasonably large cohort of 300 patients, a protein expressed by bacterial species present in 50% of the population might be identified by being linked to a clinical response in 150 patients. By comparison, a protein with a prevalence of less than 1% from a rare bacteria species would be present in only three individuals and impossible to identify in association studies.

Our goal is to translate the known healing capacity of the microbiome into drugs. Knowing that many clinically relevant factors might not be expressed by common bacteria, how did we proceed? Given the large number of microbiome proteins, we decided to apply bioinformatics filters to reduce the size of our database to something that would be appropriate for a drug discovery campaign. We also made a choice. We decided not to identify complex metabolic pathways, but instead focus on single peptides or proteins as therapeutic agents. While obviously missing smaller effectors that can deliver therapeutic efficacy, we focused on relatively simple and fully characterised peptidic entities that can be relatively easily turned into orally administered drugs.

These peptides or proteins have certain advantages. They are well tolerated and because they act on human receptors facing up the lumen of the intestine, they do not need to achieve systemic bio-distribution. Therefore relatively low amounts of the therapy can be administered. Finally, peptides are straightforward to synthesise. We refer to this class of microbiome peptides as 'endomimics.' This is because they are based on proteins secreted by gut bacteria that act like human hormones or cytokines.

These peptides were filtered systematically from our database. First, we identified proteins that were secreted by bacteria. Second, we selected small proteins and finally, we selected proteins with a rigid structure making them suitable for binding to their receptors and eligible for administration in low doses. Having started with a database of 21 million microbiome proteins, we ended up with an endomimic dataset of around 150,000.

It then took us about 18 months to synthesise a library of 20,000 endomimics. The library is used in screens conducted on cell types that are present, or have access to, the lumen side of the ileum or colon. These include dendritic cells, monocytes, macrophages and intestinal enteroendocrine cells. Only phenotypic screens were considered, meaning that at that stage in the development of our pipeline we were fully receptor agnostic.

Using a phenotypic screen on human monocytes, we identified a potent interleukin 10 (IL-10) inducing endomimic which we named EB1010. This is a cyclic, 30 amino acid secreted peptide that is structurally related to a known class of hormones. Similar to its parent hormone, EB1010 is secreted by one organ, in this case the microbiome, to act on a distant organ, in this case the intestine. The receptors for EB1010 are the same as those for the human hormones and, as expected, the endomimic shows improved resistance to proteolytic degradation. It can thus be administered orally within a gastro-protective pill. However, and different from a human hormone, EB1010 is a rare peptide found in fewer than 0.1% of individuals. It is currently in preclinical development at Enterome for the treatment of inflammatory bowel disease.

Microbiome antigens are presented to the immune system in the gut in a fundamentally different way from how other antigens are presented to the thymus. Whereas self-antigens presented to the thymus lead to the deletion of the T cells able to recognise them, the bacterial antigens presented in the gut lead to a process of T cell memorisation and tolerisation. In the event that the bacterial antigens are mimics of tumour antigens, this process leads to the generation of a pool of memory T cells in the gut with a preserved ability to recognise tumour antigens. To exploit this phenomenon and boost an immune response against tumour antigens, we have filtered for common bacterial antigens which we call oncomimics. These are similar to known human tumour antigens.

We have thus far identified 10 different oncomimics which have been combined according to different tumour types. These are being evaluated in three Phase 2 trials. These off-the-shelf treatments are being investigated with the checkpoint inhibitor nivolumab and have been administered to more than 150 cancer patients thus far. The safety of these molecules appears to be good and the immunogenicities high. We will report on potential signs of clinical efficacy soon.

Future directions

The potential for using the human microbiome as a source for new drugs is immense and we are only at the beginning of its clinical exploitation. For us, the potential will be assessed as we see the outcomes of ongoing clinical studies. On the peptide and small molecule side, we have a partnership with Takeda Pharmaceutical Co Ltd evaluating sibofimloc, a small molecule inhibitor of the virulence factor FimH which is expressed on gut bacteria. The molecule is in a Phase 2 study for patients with Crohn's disease. Separately, EB1010 is expected to enter the clinic in 2023.

On the oncomimic side, we are evaluating candidate drugs in three Phase 1/2 trials. Two more trials are expected to be launched in 2023.

We believe that endomimics can be applied to fields as diverse as intestinal inflammation, metabolic disorders, immuno-oncology, food allergies and brain diseases, through the regulation of the gut-brain axis. New technologies like the AI programme Alphafold have already allowed us to create three dimensional structures of the 150,000 endomimics in our dataset, and we have started to identify the ones which are similar to human cytokines and hormones.

These technologies represent the next generation of tools to search for candidate mimics of human hormones and cytokines. Beyond oncomimics in oncology, we believe that our antigen-mimicry platform has potential for treating autoimmune diseases as well as allergies. We are currently exploring these opportunities with partners.

References:

1. Avni O, Koren O. Molecular (Me)micry? Cell Host Microbe. 2018 May 9;23(5):576-578. doi: 10.1016/j.chom.2018.04.012. PMID: 29746828.

2. Lee KA, et al, Cross-cohort gut microbiome associations with immune checkpoint inhibitor response in advanced melanoma. Nat Med. 2022 Mar;28(3):535-544. doi: 10.1038/s41591-022-01695-5.

This article was written by Dr Christophe Bonny, chief scientific officer at Enterome SA.