proteinspotlight
> ONE MONTH, ONE PROTEIN <

Issue 245, March 2022
www.proteinspotlight.org

I'll borrow that, thank you

Vivienne Baillie Gerritsen

Who hasn't found an item and kept it – 'you never know, it might come in handy one day'? A notepad left on a bus, a paper clip ditched on the street, a biro left behind on a table. It may take weeks, months or perhaps even years, before you remember it, or happen to rediscover it, just exactly when you need it. Remember the day when, in a hurry to catch a train, you were looking for something to scribble down your phone number and you find a much-needed pen – the biro you picked up months before – in the depths of your coat pocket. It all comes down to a happy combination of chance and time, and it happens to Nature too. On rare occurrences, genomes may inadvertently pick up DNA – or genes – that did not initially belong to them. Although atypical, such an event can happen when an organism is infected by a virus, for example. The inherited DNA will probably remain silent for a long time as it is passed down the generations until, one day, it surfaces – or parts of it surface – to express a protein that turns out to be useful to the host organism. This is precisely what seems to have been the fate of PEG10. Most probably of viral origin, PEG10 is a protein capable of forming capsid-like structures while also binding to mRNA which it can then transport elsewhere.

'Allostery' by John Walther
screen print on paper with mono print (unique), 152.4 x 244cm, 2017.

Created as part of CAPSID, in collaboration with Towers Lab, UCL.
Supported by Wellcome, Arts Council England, Southwark Park Galleries, HOME Manchester and SMART Consultants.

Genomes are dynamic entities. Not only are their contents reinvigorated, so to speak, as they are passed down the generations, but they are also the seat of all kinds of mutations – many of which are of no consequence while others may turn out to be harmful to an organism, or perhaps even beneficial. Sometimes, genes of viral origin end up in our genomes – as they do in the genomes of all eukaryotes. These genes are usually muffled by the host organism and have no downstream effect. On rare occasions, however, a long time after their integration and many mutations later, they will produce a protein which proves to be useful to the host. Natural selection will do the rest, gradually coaxing the protein to become an integral part of the organism's proteome.

So genes of viral origin can be domesticated? Yes. It does not occur frequently, but it can. Initially harmful to their host (albeit muted), little by little, mutation after mutation, they can turn out to be beneficial. There are several known examples of viral gene domestication in eukaryotes. As an illustration, the vast diversity of immunoglobulin types is thought to have occurred in this way. In fact, researchers estimate that about 8% of the human genome is made up of genetic material of viral origin which, until recently, was thought to be mere genetic junk. But scientists are beginning to realise that the said 'junk' could in fact constitute the mortar upon which mammalian evolution acts. Take PEG10 which is involved in placental formation and whose occurrence in the mammalian genome – about 148 million years ago – may well mark the diversification of viviparous mammals.

Although not the case for all viruses, in order to replicate, several do integrate their genome into the host genome. Consider retroviruses whose genomes are not DNA but RNA. To multiply, they must first transcribe their RNA into DNA, which is then inserted into the host's genome. Usually, cells transcribe DNA

into RNA; retroviruses perform the reverse action, hence 'retro'. Viral genomes are small and carry the set of genes necessary for infection. Ancient sets of viral genes that have ended up in a host genome are called endogenous retroviruses or retrotransposons. Their integration can, of course, disrupt host genes and cause genetic diseases, or sometimes disturb the transcription of neighbouring genes, but they can also land outside any essential part of the host genome where they are muffled, usually by way of DNA methylation. Which is why DNA methylation is believed to be one of the most important orchestrators of mammalian evolution as it regulates the silencing – or not – of genes.

In all likelihood, PEG10 is the protein product of a domesticated ancient viral gene. How can we know? Structurally, it has many similarities with what are known as viral Gag proteins. Gag proteins assemble into capsids that form the 'body' or shells of viruses, inside which they carry their genetic material. Gag proteins have therefore two obvious roles: capsid formation and nucleic-acid binding. PEG10 seems to have retained two domains from the original viral gene, which, together, prompt capsid-like assembly of PEG10 and the binding of mRNA. PEG10, it would seem, begins by forming dimers, which are further organized to form a capsid. These capsids – with their mRNA – are then taken up by extracellular vesicles and secreted.

PEG10 seems to bind to a variety of mRNAs – among which, surprisingly, its very own! Why would PEG10 ship mRNAs? One reason would be to increase the abundance of specific mRNAs in cells, including its own. Moreover, besides delivering mRNAs to other cells, PEG10 may also serve to stabilize the mRNAs in the recipient cell, so that they are not degraded and can be translated. In this way, PEG10 has been shown to cart mRNAs into neuron cells, which – once translated – upregulate a number of genes whose products have fundamental roles in neurodevelopment. So far then, we know that PEG10 – and the faculty it has of carting mRNA – plays a part in neurodevelopment as it does in placental formation. In the same way, it seems to play a role in cellular apoptosis as well as being involved in many types of cancer, thus hinting a role in growth promotion.

The fact that PEG10 is able to carry and deliver mRNAs did not go unnoticed. Surely, then, there is a way to get PEG10 to deliver other mRNAs – ones that could be of therapeutic value for example. PEG10 could be engineered to package, export and deliver specific mRNAs to specific cells where they would be translated. This is the field of nucleic acid therapy and could complement similar approaches that currently use lipid nanoparticles as a means of delivery. Reprogramming the likes of a protein such as PEG10 to carry cargo other than that initially intended with an end to heal is both elegant and engaging. Who would have thought that, millions of years ago, one virus would turn out to be at the origin of traits as fundamental as placental formation and neural development, besides being used to taxi nucleic acids which, one day perhaps, could help to save lives.

Cross-references to UniProt

Retrotransposon-derived protein PEG10, *Homo sapiens* (Human): Q86TG7
Retrotransposon-derived protein PEG10, *Mus musculus* (Mouse): Q7TN75

References

1. Segel M., Lash B., Song J., et al.
   Mammalian retrovirus-like protein PEG10 packages its own RNA and can be pseudotyped for mRNA delivery
   Science 373:882-889(2021)
   PMID: 34413232

2. Zurowoska K., Alam A., Ganser-Pornillos B.K., Pornillos O.
   Structural evidence that MOAP1 and PEG10 are derived from retrovirus/retrotransposon Gag proteins
   Proteins 90:309-313(2022)
   PMID: 34357660

3. Kaneko-Ishino T., Ishino F.
   The role of genes domesticated from LTR retrotransposons and retroviruses in mammals
   Frontiers in Microbiology 3:262-?(2012)
   PMID: 22866050