October 19, 2018

Health & Physiology

How to transcribe the untranscribable

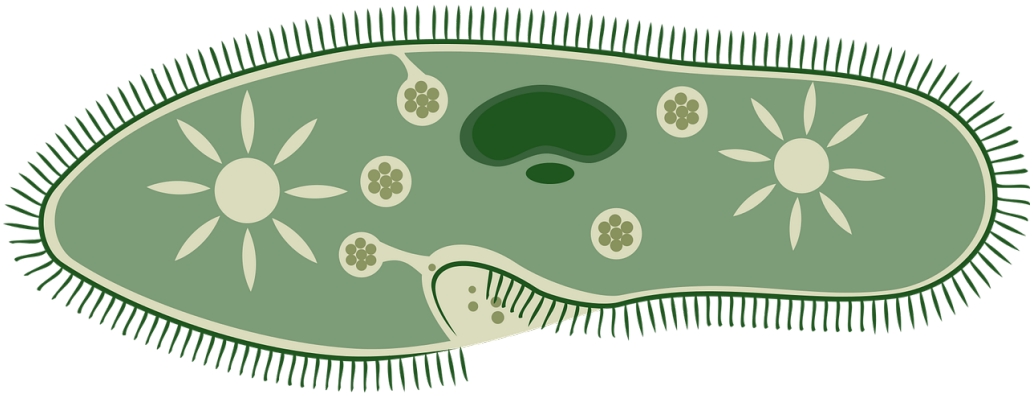
by Sarah Allen[1] | Postdoctoral Research Fellow; Mariusz Nowacki[1] | Professor

[1]: Institut für Zellbiologie, Universität Bern, Bern, Switzerland

This Break was edited by Max Caine, *Editor-in-chief* - TheScienceBreaker

ABSTRACT

Occasionally, nature comes up with a solution to a problem that is striking in its elegance. For example, generating copies of something too small to fit into the copy machinery sounds like an impossible task. Paramecium manages it in a surprisingly elegant manner.



Artist impression of a Paramecium – Image credits: Pixabay – CC0

Despite consisting only of one single cell, the microorganism *Paramecium* has an amazingly complex life cycle. It has a period of infancy and of old age, it learns, defends itself from prey, has sex, responds to different sensory cues – all things we can relate to as humans. To cope with the demands of being such a large and active cell, *Paramecium* has a huge amount of DNA: 800 copies of each of its chromosomes, compared to just two copies each in human cells.

In addition to the 800 copies of its chromosomes used for day-to-day life, *Paramecium* also keeps two sets of its DNA separate, quiet and unopened in separate nuclei. This is the *germline*, as opposed to the *soma* (Greek for 'body'). In humans, the germ cells which later become the eggs or sperm are set aside very early during embryonic development and left silent and alone. If we liken DNA to an encyclopaedia containing all the information required to build an organism, the germline is a priceless old tome kept under lock and key and only opened occasionally by specialists. Each generation, this encyclopaedia must be meticulously copied out word by word and passed on to our offspring. The DNA in all our other cells, our *soma*, is cheap copies. These copies are marked and modified - with pages crossed out, underlined or highlighted according to what is important for each particular cell. The soma is thrown out each generation when we die. Only the germline is passed on.

During evolutionary history, many unnecessary pieces of DNA have inserted themselves into the germline of most organisms, including humans. In the encyclopaedia of our germline, these pieces of

DNA are as if the scribe in charge of copying the encyclopaedia had been careless with his papers and had accidentally included a sponge cake recipe, or an instruction booklet for a flat-pack bookcase, or something else entirely. To make matters worse, this unwanted page also contains instructions for the next scribe to make extra copies of it, and stick it into several new places within the encyclopaedia.

These invasive pieces of DNA are called transposons, and exist in almost every known species. Without any preventative measures from the host organism they would eventually swamp it, diverting all its resources towards making more copies of themselves. In humans, the transposons are marked as nonsense during very early development, like crossing out a page in the encyclopaedia, but are still kept within the DNA. *Paramecium* is more ruthless, it slices the transposons out of the DNA, destroys them, and joins the ends back together. However, since the germline is kept separate and untouched, it still contains all of the copies of transposons that through evolutionary history managed to insert themselves. Only in the soma are they cut out and destroyed.

How does a cell know what is 'real' DNA and what is transposon? In both humans and *Paramecia* transposon recognition involves small pieces of RNA. RNA can be thought of as a disposable copy of a short piece of DNA, like a photocopy of a page from the encyclopaedia. These RNA copies are present during sexual development, when a new soma is made from the germline. The RNAs correspond to transposons, and so identifying the pieces of DNA to mark as irrelevant (humans) or cut out (*Paramecium*) is simple. The bits of RNA are copied from the relevant DNA pieces by RNA polymerase, an enzyme that produces matching RNA from a piece of DNA. In our analogy, RNA polymerase corresponds to the photocopier.

Because there are only two of each chromosome in human cells, a few RNA copies of each transposon is enough. But *Paramecium*, with its 800 copies of DNA, needs a specialised, secondary set of small RNA pieces that are produced from the cut out transposon pieces. The way this works is quite simple: the encyclopaedia pages that were excised using the first RNA pieces are taken and fed into the photocopier/RNA polymerase, making more and more RNA pieces to help identify all 800 copies of each transposon.

So far so simple, but after demonstrating this principle we realised that there was a logistical problem. In *Paramecium*, transposons have degraded over evolutionary time to become very small. Many of them are so tiny that they will not fit into the RNA polymerase. How can *Paramecium* make RNA copies of these tiny DNA pieces when they can't be read by the machinery?

The answer, we found, was beautifully efficient. What happens is that the tiny excised DNA pieces are stuck together, end on end, until the resulting piece is long enough that the ends can be joined to form a circle. The result is a circle of several concatenated DNA pieces which can be fed through the RNA polymerase any number of times to produce long strings of RNA. All the cell then has to do is cut the RNA up into pieces and it has as many copies as it needs of transposon-identifying RNA.

To summarise, we have discovered a new way for a cell to make small RNA pieces for identifying unwanted DNA. This is not only a remarkable proof-of-principle that cells can concatenate and circularise DNA for further use, but also demonstrates that, at least for *Paramecium*, unwanted DNA can be put to good use.