Reflections on some very nasty little things (episode 1)

Author: Jérôme Vétillard, March 24, 2020

Episode 1: What is a virus ?

Disclaimer : *I am a Digital Architect at Microsoft Consulting Services who worked as an oncologist in a past life, pioneering stem cells research and biomolecular engineering (cell sorting, genetically modified cells, cell culture …) to transplant blood stem cells. The content and the opinions expressed herein are my own personal opinions and do neither officially (or unofficially) represent my employer's views in anyway nor are intended to convey the views of Microsoft Corporation.*

Introduction

During these confinement days, I have been discussing with various fellows at Microsoft about the pandemic due to my past activities in medical research and biotechnological engineering to shed some light on their current concerns and to better understand what was going-on. Some of them encourage me to share my knowledge and reflections about the pandemic.

This is my first attempt to share some general knowledge about this pandemic. I chose to do it in English (but you can use translation platform to read it in your own language) and tried to be as didactic as I can, using analogies and as common words as I can (but sometimes you need to name a cat, a cat ^^).

Let us start with the beginning. What is a virus ?

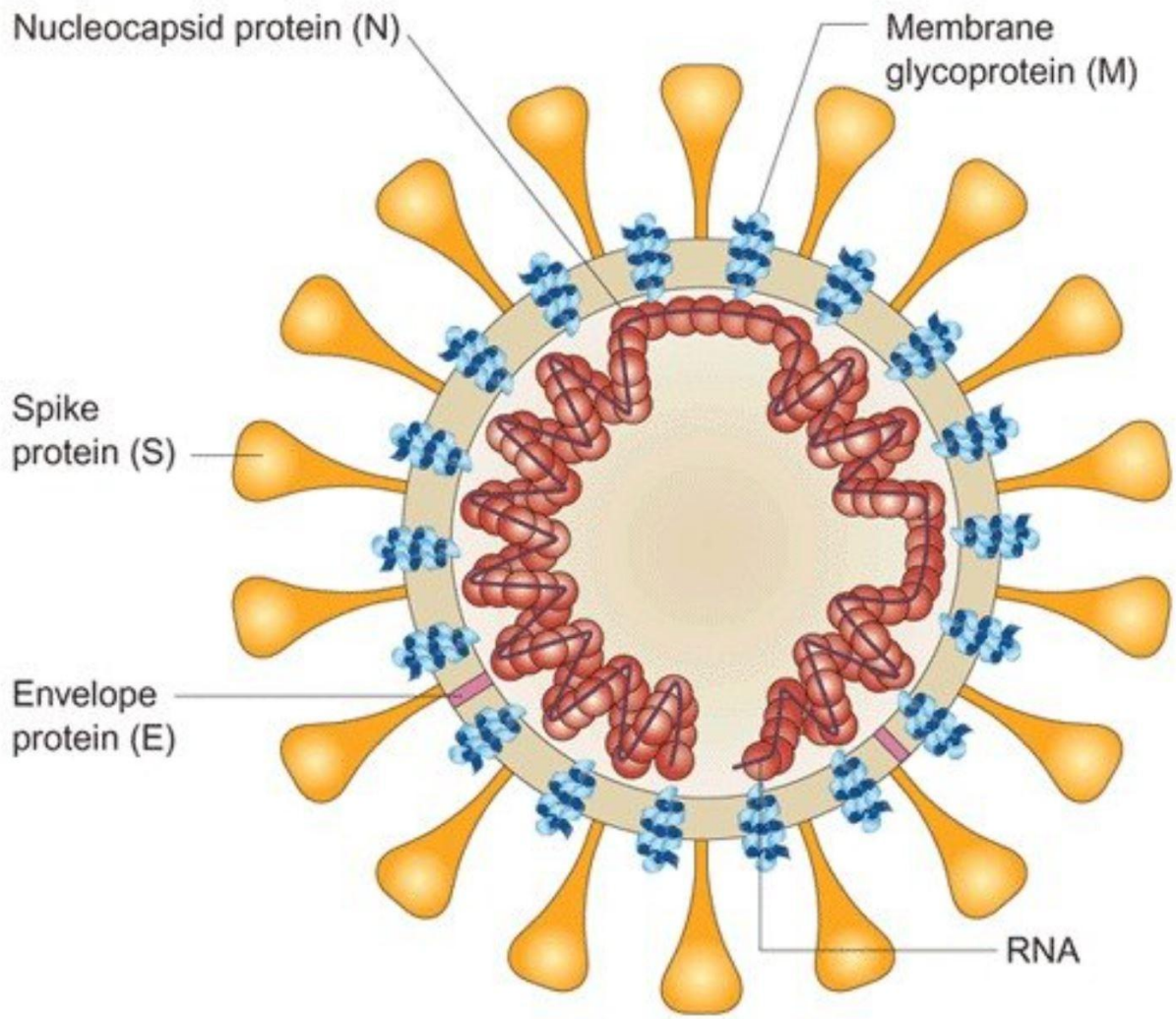
What is a virus ?

This definition relates to something that is simple and complex at the same time.

It is a biomolecular assembly with two major components:

- The envelope of the virus which protects the genetic information of the virus (inside the envelope) and exhibits glycoproteins that enable the virus to infect living cells (prokaryotic cells without a nucleus such as bacteria or eukaryotic cells with a nucleus such as the cells of your body). These surface proteins can anchor themselves onto living cells and there is a strong affinity between these virus-specific proteins, and the species (animal, plant, human being) and the kind of cells (lung, intestine, liver, kidney for SARS-COV2) the virus can infect. You can envision these surface glycoproteins like a specific molecular key, that anchors onto the to-be-infected cell surface receptor that can be envisioned as a molecular lock. And as the right key in the right lock… it opens the door, meaning the virus genetic payload is delivered within the infected cell.

- The genetic payload protected within the envelope (that is also coded by the genetic payload). To keep things simple there are two kinds of genetic information in viruses. Some viruses genetic information is DNA-based (like your own genetic information) and can be either single or double stranded DNA. They depend on a specific enzyme (DNA-Polymerase) to replicate their DNA and force the infected cells to produce new copies of the virus. Some viruses genetic information is RNA-based which is most of the time a single-stranded RNA. We have lots of very annoying little things in this category of viruses: Ebola virus, influenza virus, hepatitis C virus, coronavirus among which SARS-COV2 which is the COVID19 agent.



[source]

You can see here why it is named "corona (crowned) virus", as its spike proteins form a kind of crown around the virus envelope. You can also see the specific anchor of this virus which is the Spike Protein. The size of the virus is about 120 nm (0,00012 millimeter) which is rather small. It is a single-stranded RNA virus.

From an evolution stand point, viruses are parasites of living cells

What all viruses have in common is they are parasites. They do not have the biomolecular factory a living cell has. This is the reason why they need to infect a living cell to divert the cell's biomolecular factory and "force" it to produce copies of the virus leveraging their genetic code which codes for the whole virus structure. From a evolution perspective, it is not an egg and chicken problem: as the virus requires a functional living cell to replicate, they appeared after prokaryotes and eukaryotes. It is a kind of lean "retro"-evolution that removed everything that was not necessary to convey the genetic information, leveraging other cells' biomolecular factory to replicate this genetic information. Viruses are on the border of what could be called a "living thing", as their structure is quite simple compared to X-karyotes' structure, but their replication cycle is something very sophisticated as it relies on an infected host.

How do they infect living cells ?

Using their specific surface glycoprotein they anchor onto a specific target cell and infect it, "injecting" their genetic payload into the living cell. **Preventing this anchor (with a specific molecule that fill the "lock" before the virus can "insert" its key is one therapeutic option.** Issue there, is most of the time the "lock" is also used by regular inter cellular communication mediated through specific proteins, hence blocking the specific receptor targeted by a specific virus can also send a wrong information to the cell. As an example, HIV anchors to CD4 receptors (using its GP120 glycoprotein) that are present on the CD4-lymphocytes. Issue is CD4 is the receptor that also serves to "activate" the lymphocyte when facing a "teacher cell" that presents the antigen to the lymphocyte: blocking CD4 can prevent HIV to infect the lymphocyte, but might impair immune efficiency as well. Another way is to develop a molecule that targets the virus anchor/key and "hide" it so it cannot be "inserted"/anchor itself on the to-be-infected receptor/lock. Depending on the kind of genetic material (negative-sense, positive-sense single stranded RNA… yes! It is like when you read a book, you need to read it in the "positive" sense, not "daer" it in the "negative" sense), the virus depends on several enzymes RNA Polymerase, Reverse Transcriptase… to be able to produce a genetic information that the infected host can "read" to produce copies of the virus.

The normal process of a cell biomolecular factory

In a normal cell, the genetic information is coded into DNA. When needed, a specific gene coding for a specific protein is "transcripted": double-stranded DNA is opened at the beginning of the gene (initiation phase), a specific enzyme named RNA polymerase reads a single DNA Strand serving as a template/mold to code a single-stranded messager-RNA (elongation phase). Then when the complete gene DNA has been transcripted into RNA (termination phase), the RNA polymerase goes away, and the DNA is back to its regular double-stranded DNA (double helix well-known structure). This

happens in the cell's nucleus. The messager-RNA then leaves the cell's nucleus and migrate to the cell's cytoplasm combined with a ribosome which translates the messenger-RNA into a protein.

How does retro-virus divert the cell biomolecular factory ?

Single-stranded RNA virus might need to do the reverse process in order to divert the cell's biomolecular factory. From its single-stranded RNA, it uses a specific enzyme named "Reverse Transcriptase", to turn the virus single-stranded RNA into a double-stranded DNA. This is the reason why these viruses are also called "Retro Virus". The "Reverse Transcription" is mandatory for the virus replication, because it turns the Retro Virus genetic information/payload into something the infected host cell can use to produce copies of the virus. **Reverse Transcription phase is another therapeutic target to prevent the replication of the virus.** This is the pattern used/targeted by nucleosidic analogues such as Remdesivir from Gilead which has demonstrated efficiency on Ebola Virus to slow down the replication of the virus.

However, SARS-COV2 is not a retrovirus However, **SARS-CoV-2 is not a retrovirus**, so it does **not perform reverse transcription and does not integrate its genome into the host cell's DNA**.

Instead, SARS-CoV-2 is a **positive-sense single-stranded RNA virus**. Its genome already functions directly as messenger RNA upon entry into the cytoplasm. The host ribosomes can immediately translate viral proteins from this RNA, including the viral **RNA-dependent RNA polymerase (RdRp)**. That polymerase then synthesizes complementary RNA strands and produces new genomic RNA copies and subgenomic RNAs for structural proteins. Viral replication therefore occurs entirely in the cytoplasm, without a DNA intermediate or nuclear integration step.

Consequently, the main therapeutic enzymatic target differs from retroviruses.
Retroviruses depend on **reverse transcriptase** to convert RNA into DNA.
SARS-CoV-2 depends on **RNA-dependent RNA polymerase** to replicate RNA from RNA.

This distinction explains the mechanism of **Remdesivir**. It is **not a reverse transcriptase inhibitor**. It is a **nucleoside analogue that targets viral RNA-dependent RNA polymerase**, causing premature termination or disruption of viral RNA synthesis. Its activity reflects inhibition of RNA replication, not interference with reverse transcription.

In summary: retroviruses hijack the host by converting RNA into DNA and integrating into the genome; SARS-CoV-2 hijacks the host by using its RNA directly as mRNA and replicating through an RNA-to-RNA polymerase pathway in the cytoplasm.

Mutations rate ?

RNA-viruses have generally higher mutation rates compared to DNA-viruses.

1. RNA polymerases are less sophisticated enzymes and lack proofreading capabilities of DNA polymerase.
2. RNA viruses are often single-stranded RNA.. there's no backup on the second strand of genetic information like in double-stranded DNA. So the sole copy of genetic information is fragile and represents the last update of genetic information anytime (you cannot revert/restore it to its previous status before mutation because as I wrote, there's no "backup").
3. The Reverse Transcription reinforces the "last update", as when generating the double-stranded DNA, it somehow "confirms" (with 2 copies) that the mutation (whatever it is) is "legitimate" genetic information, somehow stabilizing the mutation during the reverse transcription phase.

This higher mutation rate makes it sometimes difficult to create a vaccine as the phenotype (the way the virus envelope looks like) is highly variable. It is like you tell your immune system to look for a tall blond with black shoes, and after its mutation, the virus is now a bald guy with plastic flip flops. Sometimes, there are also regions of the genetic payload that might be more "stable" than others and which code for more stable virus' traits. These can be targets to create vaccines.

Also stop fear about mutation into a psychotic way. I know that our collective unconscious is fed with "Zombie land", "Z-War", "28 days" movies… but the mutation is always a stochastic event (accidental) and not an intentional process. The virus undergoes the mutation "passively" and "accidentally". It does not decide to become the worst badass of the Earth and massively kill the cancerous cells of the planet.. i.e. the man kind :) Sometimes the mutation can even extinct the ability to infect human cells and the pandemic is over. Generally speaking, mutation is the engine of evolution… it creates stochastic new "possibilities/capabilities", facing a selective environment that puts a darwinian pressure on "mutants".

Recombination ?

Numerous Retroviruses are capable of genetic recombination when at least two viral genomes (genetic information payload) are present in the same host cell. This recombination capability might appear as an adaptation for coping with genome damage (remember it is a fragile single-stranded RNA genetic structure). When a region of the RNA is too damaged, as there is no "back-up" to fix it, one fixing process could be to "borrow"/"integrate" a complete RNA-region of a similar virus to replace the damaged gene(s). Such recombination may sometimes cause at outbreak of infection in humans. Let's assume that the damaged RNA area of a (corona)virus #1 (let's name it RaTG13 because it's a lovely name) was coding for the key (surface glycoprotein) targeting a specific bat of the species *Rhinolophus affinis*. Then let's assume that in the same cell

of this specific bat, there was another coronavirus #2 responsible for common cold in humans and bats (vets tell me if a bat can suffer from a common cold). Let's assume also that the genetic payload of this second coronavirus is "intact" or functional, i.e. it can infect either a bat or a human being to replicate itself. But then, alas... it is present at the same time in an " animal that is still to be identified" (it seems that the bat is the primary host, but not the host in which the recombination happened). So the coronavirus#1, which is only in the form of its damaged RNA (RNA portion coding for the anchor - the Spike protein #1 which targets bat's cells- is not working) in the cytoplasm of the infected cell, will then "borrow" the working RNA portion of coronavirus#2 coding for its anchor/key - the Spike protein #2 which targets human cells-, but will keep all other "working" portions of its RNA coding for whatever proteins/capability.. such as a possible viral pneumonia. Recombination is done... our coronavirus#1 has fixed through "accidental /stochastic" recombination its anchor/key... but now it has the anchor/key of coronavirus#2 who is able to infect human lung (intestine, kidney, liver) cells. Recombination can also happen even if the genetic information is not "damaged", but it is always a stochastic / accidental event (in Nature at least).

Conclusion

SARS-COV2 might have been the result of such recombination (see **here**).

This is the butterfly effect... that such a tiny biomolecular event on the other side of the planet, happening in the mammal cells of an animal still to be identified, gave birth to SARS-COV2, created a pandemic, is putting such a mess in our socio-economic institutions and social constructs, and eventually confronts our supposed position as the masters of the whole biosphere.

Comments

As I shall write new articles to shed some light on what's going on, do not hesitate to comment is you need to correct what's been shared here or to ask additional questions.

***Do not hesitate to share, if you think such information is worth spreading*.**

In case you miss them: [**Episode2**], [**Episode3**]