TRANSCRIPTION FACTORS TARGET TIGHTLY PACKED DNA TO INITIATE REPROGRAMMING

By Javier G. Lendinez

PhD student,
MRC Centre for Regenerative Medicine,
The University of Edinburgh

www.crm.ed.ac.uk

Centre for Regenerative Medicine

CRM is part of

THE UNIVERSITY of EDINBURGH

MRC Medical Research Council

Induced pluripotent stem (iPS) cells were first discovered by Japanese scientist Prof Shinya Yamanaka in 2006 for which he was awarded the Nobel Prize in Medicine in 2012. He found the way to make iPS cells from differentiated cells (with a defined function and morphology), using a technique now commonly referred to as "reprogramming" (see also "about stem cells" box on page 4).

Yamanaka argued that giving differentiated cells factors that normally keep **embryonic stem (ES) cells** in a pluripotent state (whereby they have the ability to give rise to any other kind of cell), might push the differentiated cells into that same state. He focused on 24 **transcription factors (TFs)** that are important for maintaining pluripotency in ES cells and observed that by introducing these factors in skin cells, they could generate cells that behaved like ES cells. He then narrowed it down by introducing combinations of smaller numbers, until he identified the four key TFs capable of reprogramming cells: **Oct4**, **Sox2**, **Klf4** and **c-Myc**.

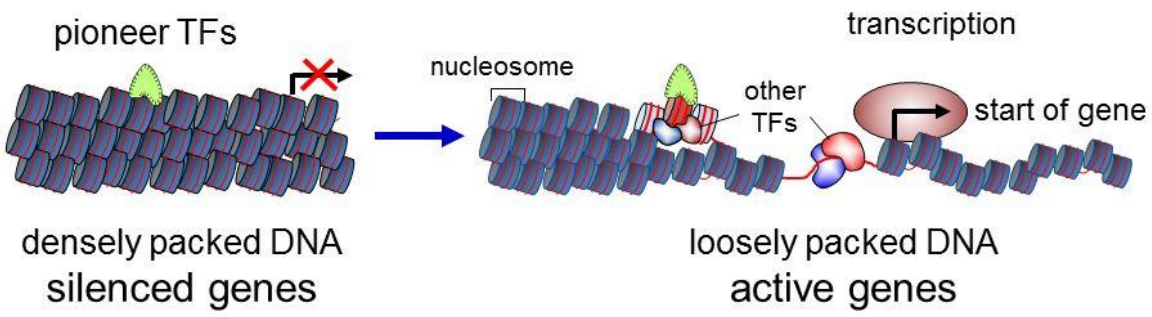
Understanding reprogramming

Since 2006 scientists worldwide are trying to better understand and refine this technique, as the process is still somehow inefficient and very variable. At this point in time these limitations make reprogrammed cells less suitable for clinical applications.

One of the key questions is to understand how the four Yamanaka TFs interact with DNA. Previous work by Kenneth S. Zaret and Abdenour Soufi at the Perelman School of Medicine (University of Pennsylvania) (Soufi et al, Cell, 2012; PMID 23159369) showed that these factors behave differently during the first stages of reprogramming and target **silenced genes** rather than just **active genes**. When genes are silenced their DNA is tightly packed and inaccessible to other TFs, whereas the SNA of active genes is uncoiled and can be read by TFs (see box "reading the DNA" on page 2). Zaret and Soufi found that three of the TFs (Oct4, Sox2 and Klf4) were able to do this independently, suggesting that they functioned as **pioneers** and were able to interact and initiate molecular changes in silenced DNA. However, c-Myc wasn't able to bind on its own to silenced DNA, and needed help of one of the pioneer TFs. Used in combination with the pioneer TFs, c-Myc notably increased the success rate of the reprogramming technique.

They also found that there were some places on the DNA to which these pioneer TFs could not bind in early stages, but were essential for iPS reprogramming. These places tended to have a specific histone modification (see box "reading the DNA"), and when the research team blocked this modification, they observed a significant increase in the efficiency of the reprogramming technique.

It became clear to them then, that to better understand how the reprogramming technique occurs, they should understand how these pioneer TFs actually bind the silenced DNA in specific regions and how they loosen it up to allow binding of other TFs.



Reading the DNA

The DNA in each of our cells is more than two meters long, all of it contained in a nucleus of a few micrometres. In order to store such amount of genetic material, the cells compact the regions that are not in use (for example the information required for other cell types). The DNA in these regions is wrapped around chromosomal proteins called **histones**, forming the **nucleosomes**, and what is called silenced DNA (histones in blue, DNA red).
In loosely packed DNA, the normal **Transcription Factors (TFs)** bind to specific DNA sequences, controlling the rate of transcription of the genetic information on the DNA. But they cannot interact with the highly packed regions of DNA and therefore special **pioneer TFs** are needed (in green) to open up the packed DNA and allow access.
During reprogramming, the Yamanaka TFs interact with the nucleosomes to loosen up the DNA and allow other TFs to act on it, and thus changing the differentiated state of the cell and transforming it into an iPS cell (silent genes → active genes).

Image by Dr. Abdenour Soufi

How do Yamanaka TFs interact with silenced DNA?

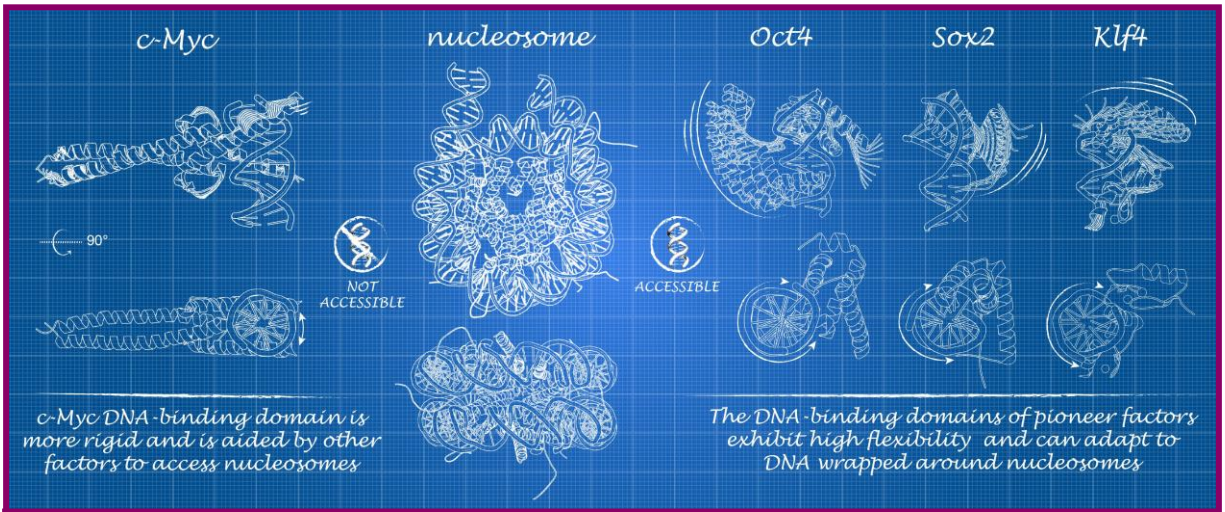
In a follow up study, done just prior to Soufi's move to the MRC Centre for Regenerative Medicine (CRM) and published in the journal Cell on 16 April 2015, Zaret and Soufi focused on the molecular interactions between the Yamanaka TFs and the **nucleosomes**, which are the basic packaging units of DNA (see box "reading the DNA"). Their aim was to understand how some TFs interact with those inaccessible areas of silenced DNA while others don't and how this relates to their ability to reprogram differentiated cells.

They discovered that **DNA binding domains (DBD)** and the way they search through the nucleosomes for their DNA binding sequences play a crucial role in activating silenced DNA. The three pioneer TFs Oct4, Sox2 and Klf4 have different DBDs with different potentials, but they all have in common the ability to recognise partial sequences of what would be their binding sequences on active DNA. In addition, the DBDs are very flexible and adapt their 3D shape to the DNA wrapped around the nucleosomes. They are also able to only recognise the part of their binding sequence that

is exposed. This allows them to search for their binding sites on the DNA that is wrapped in the nucleosomes and under normal circumstances partially inaccessible.

Moreover, they explained the two DBDs of Oct4 that can be used independently, providing extra flexibility; Sox2 seemed to have higher affinity for the bent conformation of the DNA when it is wrapped in nucleosomes; and Klf4 was shown to interact with nucleosomes although with less affinity than to free DNA because it uses only two of its three DBD for this, but with high flexibility.

The authors also found that c-Myc was not able to bind to nucleosomal DNA directly as its shape was too rigid. It needed to co-bind with another protein, like one of the pioneer TFs, before a conformational change that allowed it to bind to the wrapped DNA could take place (see box below).



The pioneer transcription factors but not c-Myc bind to nucleosomal DNA

c-Myc is too rigid, and the two "legs" of the TF would have to cover most of the DNA surface, leaving only a small fraction free to bind to histones (white arrow), which is incompatible with the nucleosome structure. The pioneer transcription factors, on the contrary, are much more flexible and are able to adapt to a much smaller fraction of the DNA, leaving the region that binds to histones free. This is compatible with the proposed mechanism in which these factors bind silenced DNA first and lay the foundations for other TFs to interact.

Image by Dr. Abdenour Soufi

Relevance of the work and future

With the knowledge acquired on how the four Yamanaka TFs interact with nucleosomes, Zaret and Soufi were able to identify a number of other pioneer factors used in reprogramming. This information will allow scientists to determine which TFs can be potential candidates for future reprogramming experiments. But this is only the first step, and the authors are already working to understand how these pioneer TFs open up the chromatin and allow the formation of the large protein complexes which facilitate gene activity. Understanding the entire process will allow scientists to control the reprogramming process in a much more refined way. Going forward,

Abdenour Soufi is also planning to use this information to design novel factors with improved reprogramming activity at CRM.

About Stem Cells?

Stem cells were first identified in the 1960s to describe a type of cell with the ability to give rise to a number of other cell types, in a process called differentiation. In addition to this, stem cells also have the capacity to self-renew, the process by which they can produce exact copies of themselves. This is a fundamental system to maintain our body and has also become an important research tool with the potential to study numerous diseases and, ultimately, generate specific cells and tissues for treatments

One way to think about stem cells is to divide them into three categories:
1. Adult or tissue stem cells: found only in specific areas of the body;
2. Embryonic stem (ES) cells: grown in the laboratory from the early embryo;
3. Induced pluripotent stem (iPS) cells, or 'reprogrammed' stem cells: similar to embryonic stem cells but made from adult specialised cells using a laboratory technique discovered by Japanese scientist Prof Shinya Yamanaka in 2006.

Tissue or adult stem cells are multipotent and sometimes even unipotent, they can only give rise to limited, more tissue-specific cells. For example, multipotent blood stem cells can only make the specialised cells found in your blood and unipotent germline stem cells in males can only make one type of cell, sperm cells. ES cells are pluripotent, meaning they have the capacity to produce every type of cell in the body. The 'reprogrammed' iPS cells behave identically to the pluripotent ES cells and share the capacity to make every type of cell in the body. These cells, unlike ES cells, do not require the use of embryos.

Why the need for iPS cells?

ES cells allowed scientists to study the cell differentiation process and make differentiated cells to compare diseased and healthy conditions with the end goal to make a large amount of cells to transplant into patients. However, they have limitations, such as their source. They are obtained from fertilized eggs and therefore the ethical implications can be controversial. Moreover, patient specific cells cannot be produced from ES cells yet.

iPS cells, on the contrary, can be created from somatic (differentiated i.e. skin cells) cells and have the same capabilities of self-renewal and differentiation as ES cells. They can be used to study diseases and cell differentiation processes as well. But their mayor advantage is that they can be made from patients´ own cells and reprogrammed in the lab into iPS cells to then make the desired healthy cells for a personalised medical treatment.

Publication details

Abdenour Soufi, Meilin Fernandez García, Artur Jaroszewicz, Nebiyu Osman, Matteo Pelegrini and Kenneth S. Zaret, 2015. Pioneer Transcription Factors Target Partial DNA Motifs on Nucleosomes to Initiate Reprogramming. Cell 161, 1–14. http://dx.doi.org/10.1016/j.cell.2015.03.017.

This study, published online on 16 April 2015, was done at the Perelman School of Medicine (University of Pennsylvania) prior to Abdenour Soufi's move to the MRC Centre for Regenerative Medicine (University of Edinburgh).