

Cellular barcoding: giving names to every stem cell

There are many different cell types in the blood, all of different functions and shapes. All these cells are made by the stem cells by a process of lineage commitment and differentiation. How can we monitor this process of blood formation? This has been a difficult question to answer and researchers have tried many tricks. Cellular barcoding is one of the recent tools, potentially allowing any living cell of the body acquire a unique DNA tag, or barcode, in other words obtaining a name or address, so any progeny of each cell can be traced.

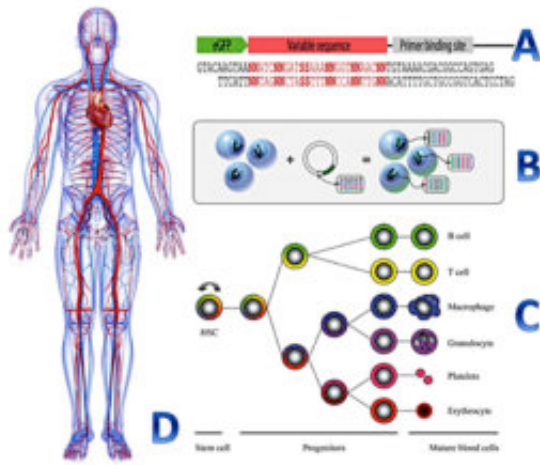


Fig. 1. The cellular barcoding utilises random or semi-random DNA sequence (A), which is ligated into the retro- or lenti-viral vector and used to modify the genome of the hematopoietic cell (B). This will allow spreading the barcode from the hematopoietic stem cell along all hematopoietic lineages (C). Theoretically, such barcoded hematopoietic system could be transplanted into the human or other animal body (D) and used for live monitoring of the hematopoiesis.

For barcoding, we can design and synthesize short unique DNA fragments, which can be inserted into the special DNA vector, capable getting into the living cell and integrating into each cell's genome. This will also insert a unique DNA tag into the genome of every infected cell. What happens next? If it was a fully functional differentiated cell, it would live and do its job for a while, but eventually die out and be replaced by other cells. If it is a stem cell, it will multiply itself for a considerable duration of time and also generate mature cells, so after a while we will be able to detect in the blood a considerable number of different cells carrying the same barcode. The barcode will tell us that all those cells come from the same ancestor cell. This will allow us to count retrospectively the number of stem cells contributing to blood regeneration. It will also tell us what categories of cells are made from the same stem cells, and probably many other things.

Like any other method barcoding has its limitations and caveats. Ideally, we would like to have

such barcodes, which would be unique for every cell we label, and we would like to make sure we could read those barcodes without much trouble and confusion. There are many different names, codes and coding systems around us, some are better than others. Human names for instance are no so good for unique labeling of individuals, such system is poor regarding uniqueness and redundancy. Not for nothing each of us has also social security number, passport number or other ID with a number. Unlike human names those numbers are really unique for each person. The second condition for a good code is good degree of difference between any pair of name codes used. From the past we know a problem of the wrong telephone calls. Why? Because they were made in a simple numeric order, so one digit error could cause a call to the wrong number. License car plates are a bit better, they exclude too similar characters and follow some other rules reducing the risk of error. Finally, digital telecommunication systems (in our laptop or TV for instance) use robust coding algorithms minimizing and correcting errors, when it happens.

For barcoding of cells in a fish, mouse or human body we use the combinations of four bases of DNA: A,T,C and G. Currently for barcoding, scientists massively use random stretches of DNA (although the fashion might change). Just like in examples above we can use these DNA barcodes with less or more security and robustness. Generally, we should use the knowledge of the coding theories; try to adapt them to this field of biology. When implemented correctly we have to deal with a limited number of parameters like numbers of possible barcode combinations to the number of cells to barcode (actually the code should use many more combinations than the number of cells). We should try some tricks to make barcodes unique and easily recognized. We also have to do the best to avoid confusions or misunderstandings in reading those barcodes using our current sequencing techniques.

Although the theory sounds relatively easy, the implementation is still challenging. Consequently, there is still a wide range of opportunities to improve our cellular barcoding tools.

Leonid V. Bystrykh

Senior scientist at ERIBA, UMCG, University of Groningen, The Netherlands

Publication

[Clonal Analysis of Cells with Cellular Barcoding: When Numbers and Sizes Matter.](#)

Bystrykh LV, Belderbos ME

Methods Mol Biol. 2016