

IRIS A<sub>per</sub>TO



UNIVERSITÀ  
DEGLI STUDI  
DI TORINO

**This is the author's final version of the contribution published as:**

Donati G.

The niche in single-cell technologies.

Immunol Cell Biol. 2016 Mar;94(3):250-5. doi: 10.1038/icb.2015.107. Epub 2015 Dec 22.

**The publisher's version is available at:**

<http://onlinelibrary.wiley.com/wol1/doi/10.1038/icb.2015.107/full>

**When citing, please refer to the published version.**

**Link to this full text:**

<http://hdl.handle.net/2318/1572447>

This full text was downloaded from iris-AperTO: <https://iris.unito.it/>

# Niches in Single Cell technologies

Giacomo Donati<sup>1</sup>

<sup>1</sup>Centre for Stem Cells and Regenerative Medicine, King's College London, 28<sup>th</sup> Floor, Tower Wing, Guy's Campus, London SE1 9RT, UK

## Abstract

The niche is the microenvironment in which each cell exists and is able to keep its own peculiar characteristics. The importance of the niche has been intensively studied especially in the context of stem cells as the niche is responsible for both the stemness maintenance and the activation of differentiation. In the past few years a variety of single cell technologies have shed light on the extraordinary variability that characterizes different stem cell populations both *in vitro* and *in vivo*, but in most of the cases the positional information has been lost. Recent developments of new technologies aim to integrate both the transcriptomic profiling of cells and the spatial location. In this review I'll discuss the state of the art of these technologies and the future development that will be of vital importance in the study of stem cell populations.

## **Introduction**

Removing a cell from its niche it is like removing a word from the sentence it belongs to. This might sound as a too simplistic view, but for some words with more than one meaning, try to understand the essence is impossible without a precise context (figure 1). The same is true when, instead of a single world, we attempt to analyse single cells transcriptomes or protein expression levels, after taking them out from their original environment.

The microenvironment where each cell resides is called niche. The concept of niche has been introduced more than four decades ago for the human haematopoietic system (ref Trenton, J. J. in Regulation of Hematopoietic Stem Cells (ed. Gordon, A. S.) 161–185 (Appleton–Century–Crofts, New York, 1970). The crude definition of niche is the area of a tissue that provides to the cells a specific microenvironment. For stem cells, that are directly responsible for the generation and maintenance of diverse tissues in our bodies, the idea of niche is even more important as it is the location where they are present in an undifferentiated and self-renewable state. Cells adjacent to each other provide an abundance of diverse stimuli to each other, that is responsible for maintaining the stem-cell fate and prevents their differentiation.

Other than the cell-to-cell interaction the niche is characterized by a complex interplay of other different factors as the extracellular matrix components, the oxygen tension, the concentration of cytokines, chemokines and growth factors and the physicochemical status of the environment including the pH, ionic strength (e.g.  $\text{Ca}^{2+}$  concentration) and metabolites, like ATP, are also important (ref).

In the last few years we observed a fast development of a plethora of single cell technologies that span from the proteomic analysis of single cell (ref) to the genomic (ref) and to the global transcriptomic analysis.

Even if those techniques have realized one of the biggest scientific dreams of obtaining unbiased and global information of cells without any previous knowledge, they are mostly appropriate when studying cells which original spatial location is not crucial, for instance blood circulating cells. When scientists attempt to deeply analyse the behaviour of a cell in its own niche, the ability to retain the positional information is essential. This is true for example for neurons in the different regions of the brain (ref) as well as for distinct compartment of epidermal stem cells (ref).

Some progresses have recently been made toward this direction and in this review I will describe the state of the art of these technologies in the study of single cells in their own original context.

### **The cellular niche in the single cell era**

Single cell sequencing has become widely used in the last few years to analyse the global transcriptomic profile of different cell types as hematopoietic stem cells (Ref), single neurons (ref), epithelial cells (lung paper) and cancer cells (ref).

Despite different protocols of RNA extraction, reverse transcription and library preparation exist, the first step of the sample preparation always consists in the tissue dissociation. In this way all the cellular spatial information is completely lost.

In the study of primary cells from specific tissues, the ability to retain spatial information on each single cell is important as many different cell types are adjacent to each and the architecture reflects the cell function. This is particularly relevant, for example, in the brain, in the skin and ...

The isolation of single cells from tissue sections using either laser-capture microdissection <sup>1</sup>, FACS <sup>2,3</sup> or a fine controlled pipetting system<sup>4</sup> can somehow partially resolve these limitations, but all these techniques have a low throughput of cell number and spatial resolutions.

Due to the massive development of single cell sequencing in the past few years, scientists started to investigate the possibility to couple the whole-genome gene expression profiling with the positional information for each cell.

At the moment, few methods for high-throughput, spatially resolved single-cell RNA-seq have been developed using both in vivo marking of selected cells from predefined spatial coordinates, or in situ amplification of cellular transcriptomes on tissue sections.

Methods available so far can be divided in two simple categories: the ones that try to keep the spatial information before the analysis and the ones that, a posteriori, try to infer the original position from the data obtained (Table 1).

### **Top-down and bottom-up experimental strategies**

Recently, alternative approaches for profiling the transcriptomes of spatially referenced cells have been proposed. Transcriptome in vivo analysis (TIVA) allows individual cells in a tissue to be visualized by fluorescent labels and then

sequenced after capture. This approach is perfect when looking at few, specific cells, while, when looking at a big cell population some steps in the protocol might be limiting. This is true for example for the manual photoactivation and picking steps and also the number of tags that can be employed simultaneously. By contrast, fluorescent in situ RNA-seq (FISSEQ)<sup>5</sup> sequences individual cells in situ (i.e., directly on cells or tissue mounted on a coverslip). Such an approach can, in theory, be broadly applied, although the practical challenges are substantial. Moreover, only cells that are relatively close to one another can be assayed in

One example that belongs to the second category is the computational strategy presented by Satija and c-workers<sup>5</sup> named Seurat. The computational method is named after the famous pointillist painter George Seurat, and it well represents the idea of assigning the single isolated point to the figure it belongs to. The authors applied Seurat to map the RNA-seq data from 851 cells of the zebrafish embryo (*Danio rerio*) in the late blastula stage to their own original position. This particular stage is well studied by in situ hybridization for known drivers of embryonic patterning and gastrulation (ref zebrafish at the bottom). After dividing the embryo in 128 bins, containing between 40 and 120 cells each, they used published bright-field images of *in situ* hybridizations from 47 genes in order to generate a reference spatial map of the embryos. Seurat finally maps cells to their location by comparing the expression level of a gene measured by single-cell RNA-seq to its expression level in a three-dimensional tissue measured by *in situ* hybridization. This approach is able not only to better

characterize, from the transcriptomic point of view, known subpopulations of cells but also to identify new rare subpopulations within the tissue under study.

A similar approach has been implemented by Achim and co-workers <sup>6</sup>. They proposed integrated approach that combines previously generated in situ hybridization (ISH)-based gene expression atlases for 98 genes with 155 unbiased single-cell transcriptomics. As a model they employed the developing brain of a marine annelid, *P. dumerilii*, a widely used model organism for the study of molecular developmental processes of bilaterian animals <sup>7</sup>. First of all, they binarized the gene expression atlas was binarized, resulting in a matrix of  $n$  positions that each comprise presence and absence values for  $m$  genes. With a 3-steps computational strategy they mapped 50% of the cells with high confidence, 31% with medium confidence and 10% with low confidence. Because ISH atlases are available for many species and developmental stages this approach is also applicable to a great variety of different models as for example mouse and human brain and mouse and chicken developing embryos <sup>6</sup>.

### **FISSEQ, the evolution of fluorescence in situ hybridization**

Single-molecule fluorescence in situ hybridization (FISH) has been extensively adopted for the quantification of the abundance of different transcript at single-cell resolution within the context of a tissue of interest <sup>8</sup>. With RNA FISH individual RNA molecules can be absolutely quantified in a very accurate way through a microscope, keeping the information on RNA location down to the subcellular level. Despite allowing the quantification of the gene of interest in numerous cells, this technique can be applied to a very limited number of genes at a time <sup>9</sup>.

In order to overcome these limitations Lee and colleagues <sup>12</sup> develop a new, non-invasive method named FISSEQ (Fluorescent in situ sequencing), based on earlier work by Church et al <sup>13</sup> of fluorescent in situ sequencing on polymerase colonies. Starting from fixed and permeabilized cells or tissues they perform in situ reverse transcription employing random hexamers tailed with a sequencing adaptor. The resulting cDNA is then circularised and amplified via RCA to generate DNA nanoballs. Each amplicon contains numerous tandem copies of the cDNA template and adapter sequence allowing the sequencing with SOLiD (sequencing by oligonucleotide ligation and detection) for up to 27bp reads with 99.4% accuracy.

They employed FISSEQ in human human iPS cells, human primary fibroblasts in homeostatic conditions and after simulating a response to injury and in the whole embryos, too. FISSEQ correlates well with standard RNA-seq (Pearson's *r* ranged from 0.52 to 0.69) and with gene expression array as well (Pearson's *r* 0.73). Despite it generates far fewer reads respect to RNA-seq it mainly detects genes characterizing cell type and function. They demonstrated that nuclear RNA was twice as likely to be non-coding, and antisense mRNA was almost twice as likely to be nuclear. At this stage of development, FISSEQ is not able to account for all the RNA molecules in a cell, but detect a well-localized subsample of the transcriptome, and the comparisons with RNA-seq and microarray data suggest that FISSEQ may miss lower-abundance transcripts.

Despite this technique relies on a very sophisticated image analysis, including robust identification and validation of signals within the image and, on a technical level, a major limitation of the current protocol is the lack of ribosomal



RNA depletion and the lack of information on biases. Similarly to next generation sequencing, further development are expected, as the increase in read length, sequencing depth and coverage, and progresses in the library preparation. Such advances may lead to improved stratification of diseased tissues in clinical medicine.

### **TIVA alias in vivo transcriptomic analysis**

At the same time, another team of scientists led by Dmochowski and Eberwine<sup>10</sup> engineered a multifunctional photoactivatable mRNA capture molecule, named TIVA (Transcriptome in vivo analysis), to perform transcriptome analysis of individually selected cells in intact tissue. The TIVA tag is a multifunctional mRNA-capture molecule and has many features: is permeable to the cell membrane, has a photocleavable linker, a pair of fluorophores (Cy3 and Cy5), a poly(U) oligonucleotide and a biotin tag. The fluorophores provide a fluorescence resonance energy transfer (FRET) signal allowing real-time monitoring of cellular uptake as well as uncaging.

After entering the cells the TIVA tag dissociates from the peptide that will be trapped in the cell. The first step of photoactivation will be performed to a specific cell only, allowing the uncaging of the poly(u) mRNA-capturing that will anneal to a poly(A) tail of cellular mRNA. The photoactivation can be visualized by fluorescence resonance energy transfer through exploiting the Cy3 and Cy5 molecules on the tag. Finally the mRNAs bound to the TIVA tag can be purified with streptavidin beads and sequenced by any RNA-seq method. Although all the cells in the tissue will contain TIVA tags, these tags will not hybridize to cellular

mRNAs as their poly(U) oligonucleotides are normally hidden. In their manuscript and co-workers employed TIVA-tag in isolated single neurons in culture and in mouse and human tissue *in vivo*, showing that hippocampal neurons in live tissue expressed fewer genes but had more bimodally expressed genes than the same neurons in culture, as first described by Shalek et al <sup>11</sup>. This observation suggested a crucial role for the cell environment in the fine modulation of gene expression in the cells.

TIVA will have potentially a great impact on the single cell research field, primarily because it allows a high resolution analysis of single neighboring cells using commonly available tools, making its application easy to

This was the first method described for the study of global expression profiling in cells within intact tissues with a crucial improvement respect to methods that involve either Laser capture microdissection or pipetting that can damage the cell as well as trim it.

Finally there are two more methods of targeted in situ RNA-seq, both of them have been developed at SciLifeLabs. One is the Padlock RCA-seq and the other is the Spatial Transcriptomics.

A revolutionary method, named spatial transcriptomic, has been developed by Joakim Lundeberg and Patrik Stahl. In this protocol, a freshly frozen tissue section is first imaged in order to retrieve histological information and then

attached onto the chip. The chip contains an array of distinguishable Poly-T tailed capture probes with unique 5' barcodes. This chip is five by five millimeters and is divided up into 135,000 features. After fixation and permeabilization, the cellular mRNA diffuses directly from the tissue section onto the microarray to hybridize to the oligo-dT, followed by in situ cDNA synthesis. The cDNA can now be sequenced with any sequencing method, so the read depth can be decided according to the aim of the study. The chip analyses 135,000 cells at a time and, at the moment, the spot density is the only limiting factor. With future development it can also allow up to 2 million cells x run. This protocol that has now been applied to mouse brain only, can potentially be applied to a range of other sample types, especially as improvements will be made to resolution. This technique is now commercialised by Spatial Transcriptomics, who are now offering early access to few, selected projects, but it will probably be available to a broader research community very soon. Overall, this method perfectly combines a complete transcriptome profiling at single cell resolution with conserved histological information. Given this premises, Spatial Transcriptomic have the potential to become the strategy of choice for many single-cell transcriptomics applications where the spatial information is crucial.

1. Keays KM, Owens GP, Ritchie AM, Gilden DH, Burgoon MP. Laser capture microdissection and single-cell RT-PCR without RNA purification. *J Immunol Methods* 2005; **302**(1-2): 90-8.
2. Dalerba P, Kalisky T, Sahoo D, Rajendran PS, Rothenberg ME, Leyrat AA *et al.* Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat Biotechnol* 2011; **29**(12): 1120-7.
3. Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J *et al.* Tumour evolution inferred by single-cell sequencing. *Nature* 2011; **472**(7341): 90-4.
4. Li Y, Xu X, Song L, Hou Y, Li Z, Tsang S *et al.* Single-cell sequencing analysis characterizes common and cell-lineage-specific mutations in a muscle-invasive bladder cancer. *Gigascience* 2012; **1**(1): 12.
5. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol* 2015; **33**(5): 495-502.
6. Achim K, Pettit JB, Saraiva LR, Gavriouchkina D, Larsson T, Arendt D *et al.* High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nat Biotechnol* 2015; **33**(5): 503-9.
7. Fischer AH, Henrich T, Arendt D. The normal development of *Platynereis dumerilii* (Nereididae, Annelida). *Front Zool* 2010; **7**: 31.
8. Levsky JM, Singer RH. Fluorescence in situ hybridization: past, present and future. *J Cell Sci* 2003; **116**(Pt 14): 2833-8.
9. Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods* 2008; **5**(10): 877-9.
10. Lovatt D, Ruble BK, Lee J, Dueck H, Kim TK, Fisher S *et al.* Transcriptome in vivo analysis (TIVA) of spatially defined single cells in live tissue. *Nat Methods* 2014; **11**(2): 190-6.

11. Shalek AK, Satija R, Adiconis X, Gertner RS, Gaublomme JT, Raychowdhury R *et al.* Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature* 2013; **498**(7453): 236-40.
12. Lee JH, Daugharthy ER, Scheiman J, Kalhor R, Yang JL, Ferrante TC *et al.* Highly multiplexed subcellular RNA sequencing in situ. *Science* 2014; **343**(6177): 1360-3.
13. Mitra RD, Shendure J, Olejnik J, Edyta-Krzymanska-Olejnik, Church GM. Fluorescent in situ sequencing on polymerase colonies. *Anal Biochem* 2003; **320**(1): 55-65.