It is well known that genetically identical cells can display a high variability in their gene expression profiles. This phenomenon has a profound impact on a myriad of cellular processes, ranging from differentiation to signaling and drug resistance. Large efforts are therefore currently devoted to understand the mechanistic causes and functional consequences of cell-to-cell variability, often called noise (Lionnet and Singer, 2012).

There are multiple sources of cell-to-cell variability: cell cycle stage, circadian clock, metastable epigenetic states, fluctuations in the concentration of regulatory factors, inhomogeneous microenvironments, or the stochastic nature of the molecular steps involved in gene expression. These factors are often hard to separate experimentally because they might be unknown a priori and are often challenging to control: they can range from intracellular concentrations of upstream factors to cell shape or extracellular context.

As most genetic circuits involve a vast number of genes, it has proven extremely useful to study genome-wide transcriptomes in order to understand the determinants of gene expression variability. The first applications of microarray profiling to single cells were demonstrated a decade ago (Klein et al., 2002), and RNAseq-based methods have recently contributed to increase the assay sensitivity. As a result, these techniques can now map expression data onto the full genome and identify splicing variants. However, scaling the expression spectrum (the dynamic range of the FISH technique is slightly lower for highly expressed transcripts) makes it necessary to measure inaccessible to cDNA-based transcriptomics. One important standard for the developing field of single-cell transcriptomics is that for most genes, at least 1000 individual cells should be analyzed to recapitulate the mRNA copy number distribution in a reproducible fashion. This finding will constitute an important standard for the developing field of single-cell transcriptomics.

Single-molecule mRNA fluorescent in situ hybridization (smFISH) is the main approach complementing sequencing and microarray-based techniques (Itzkovitz and van Oudenaarden, 2011). It consists in hybridizing multiple fluorescent DNA probes to a given mRNA species in a fixed biological sample. Individual mRNA molecules appear as individual spots and can be counted using dedicated algorithms. The technique has the advantage of preserving the integrity of the sample, and thus allows capturing a wealth of parameters (e.g., cell shape and location, mRNA spatial distribution (Chou et al., 2013) or the expression pattern of a protein of interest) that are usually lost in techniques based on cDNA libraries. The main limit of smFISH is its modest throughput: the number of genes one can simultaneously image is limited by the number of spectrally separable fluorescent dyes (~5). Barcoding approaches have increased this number to ~30 (Lubeck and Cai, 2012), but these numbers remain exceedingly low compared to the tens of thousands genes composing the human genome. Measuring larger number of genes by smFISH has so far only been possible using artificially labeled reporters in bacteria (Taniguchi et al., 2010).

In a recent article, Battich et al. (2013) have demonstrated an automated pipeline for smFISH that allowed them to interrogate separately ~1000 endogenous genes, collecting data from ~11 000 individual cells for each gene. This experimental tour de force relies on using a variant of smFISH termed ‘branched DNA smFISH’ (bDNA smFISH). Instead of directly labeling the transcripts with fluorescently labeled probes, the technology uses a combination of primary, secondary and tertiary probes that hybridize together in order to label each target site on a given mRNA with tens of fluorescent labels (Figure 1B). As a result of the increased signal, fluorescence images could be acquired faster than with traditional smFISH, using only a low-magnification microscope objective. This allowed scanning a cell population faster, resulting in an increase in the throughput of the technique. The sensitivity of the bDNA smFISH rivals that of RNAseq over most of the expression spectrum (the dynamic range of the FISH technique is slightly lower for highly expressed transcripts). Using their unprecedentedly large data sets, the authors tested the statistical requirements of single-cell mRNA counting. They found that for most genes, at least 1000 individual cells should be analyzed to recapitulate the mRNA copy number distribution in a reproducible fashion. This finding will constitute an important standard for the developing field of single-cell transcriptomics.

The main advantage of the approach lies in its image-based nature. Using an integrated image analysis pipeline, the authors were able to extract a battery of spatially resolved measurements inaccessible to cDNA-based transcriptomics techniques. This information is crucial to investigate determinants of cell-to-cell variability; for instance, as biochemical reactions are dependent on factors concentrations rather than numbers, simply knowing the cell volume is important to normalize copy number fluctuations. Furthermore, the authors found that mRNAs sharing statistical and
spatiotemporal expression patterns were likely to encode interacting proteins. This finding demonstrates the important role of mRNA (co)localization in gene expression and may suggest a mechanism explaining why functionally related proteins display correlated expression levels, whereas their respective mRNAs levels are essentially uncorrelated (Gandhi et al., 2011).

Image-based, multivariate approaches will play a crucial role in understanding the determinants of cell identity and variability because they are able to collect single-cell transcriptomes along with information about the respective environment, morphology and eventually function of each cell. As smFISH approaches high throughput, these advantages will make it a major tool for understanding the regulation, function and dysfunction of gene expression heterogeneity.

Conflict of interest

The author declares that he has no conflict of interest.

References

Battich N, Stoeger T, Pelkmans L (2013) Image-based transcriptomics in thousands of single human cells at single-molecule resolution. Nat Methods 10: 1127–1133
Chou YY, Heaton NS, Gao Q, Palese P, Singer R, Lionnet T (2013) Colocalization of different influenza viral RNA segments in the cytoplasm before viral budding as shown by single-molecule sensitivity FISH analysis. PLoS Pathog 9: e1003358
Gandhi SJ, Zenklusen D, Lionnet T, Singer RH (2011) Transcription of functionally related constitutive genes is not coordinated. Nat Struct Mol Biol 18: 27–34
Itzkovitz S, van Oudenaarden A (2011) Validating transcripts with probes and imaging technology. Nat Methods 8: S12–S19
Klein CA, Seidl S, Petal-Dutton K, Olifer N, Geigl JB, Schmidt-Kittler O, Wendler N, Passlick B, Huber RM, Schlimok G, Baueerle PA, Riethmüller G (2002) Combined transcriptome and genome analysis of single micrometastatic cells. Nat Biotechnol 20: 387–392
Lionnet T, Singer RH (2012) Transcription goes digital. EMBO Rep 13: 313–321
Lubeck E, Cai L Single-cell systems biology by super-resolution imaging and combinatorial labeling. Nat Methods 9: 743–748 (2012)
Shalek AK, Satija R, Adiconis X, Gertner RS, Gaublomme JT, Raychowdhury R, Schwartz S, Yosef N, Malboeuf C, Lu D, Trombetta JJ, Gennert D, Gnirke A, Goren A, Hacohen N, Levin JZ, Park H, Regev A (2013) Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature 498: 236–240
Taniguchi Y, Choi PJ, Li GW, Chen H, Babu M, Hearn J, Emili A, Xie XS (2010) Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. Science 329: 533–538
Wills QF, Livak KJ, Tipping AJ, Enver T, Goldson AJ, Sexton DW, Holmes C (2013) Single-cell gene expression analysis reveals genetic associations masked in whole-tissue experiments. Nat Biotechnol 31: 748–752

Molecular Systems Biology is an open-access journal published by the European Molecular Biology Organization and Nature Publishing Group. This work is licensed under a Creative Commons Attribution 3.0 Unported Licence. To view a copy of this licence visit http://creativecommons.org/licenses/by/3.0/.