➜ TECHNOLOGY

DNA as a biological lens

> reconstruction of a virtual image of the relative positions of all the interrogated mRNAs

New research in *Cell* describes the technique of 'DNA microscopy', whereby images of the spatial distribution of cell biomolecules are generated without direct visualization, by using DNA sequencing to infer the relative proximities of mRNAs based on incorporated DNA tags.

In this multi-step biochemical method, mRNAs are first tagged in fixed human cells through cDNA synthesis using DNA primers that anneal to chosen mRNA targets and also contain a unique molecular identifier (UMI) sequence. The UMI is a random sequence tag that uniquely labels each separate copy of an mRNA of interest. These tagged cDNAs are then amplified by PCR using overlap extension primers for two purposes: to generate many copies of each tagged cDNA that can then diffuse from the site of amplification; and to allow hybridization and concatenation of two different tagged cDNAs that occupy the same physical location. Paired-end sequencing of the concatenated products reveals several types of information: the identity of the two colocalized mRNAs (based on the native cDNA sequence); which starting molecule of those mRNAs was involved (based on the two cDNA UMIs); and the unique concatenation event based

on random sequence contributed by the overlap extension primers at the concatenation junction.

The DNA microscopy system has numerous informatics challenges, such as the complexity of the information encoded in the concatenated products and the need to model cDNA diffusion properties and any biases and errors in the PCR and sequencing steps. Therefore, much of the study was dedicated to devising and optimizing the data analysis strategies to achieve a meaningful virtual image of the interrogated mRNAs.

Algorithms were optimized using simple proof-of-principle cell systems, such as mixtures of two human cell lines that expressed either green fluorescent protein (GFP) or red fluorescent protein (RFP). Analysing *GFP* and *RFP* mRNAs, as well as *GAPDH* and *ACTB* housekeeping mRNAs, ensured that the algorithms could correctly identify *GFP* and *RFP* expression as being mutually exclusive but *GAPDH* and *ACTB* as being co-expressed. Moreover, the algorithms could determine the extent to which the computationally generated images matched the regular fluorescence microscopy images collected before DNA microscopy.

A key feature of the analysis is that the number of concatenation events between two cDNA molecules allows estimates of the original distances between their two source mRNAs. Mapping these pairwise distances across the entire data set enables reconstruction of a virtual image of the relative positions of all the interrogated mRNAs.

Additional refinements to the algorithms included optimizing the weighting of local analyses restricted to subsets of closely interacting cDNAs versus global analyses across the entire interaction network of mRNAs. Furthermore, image segmentation guided the output map into the known discrete arrangement of mRNAs into individual cells.

As a demonstration that DNA microscopy can be scaled up to more genes to reveal relevant biology, the authors interrogated up to 20 mRNAs in the *GFP*- versus *RFP*-expressing cell lines and recapitulated known expression level differences from bulk RNA sequencing (RNA-seq) analyses on the individual cell lines.

DNA microscopy thus shows promising potential as an emerging virtual imaging modality. Key avenues for future development include exploring how many different transcripts can be studied simultaneously, the subcellular resolution achievable and its applicability to solid tissue sections. In addition, it would be of interest to determine whether incorporating a grid of positioned DNA molecules into the experimental design could help to anchor relative mRNA distances into accurate absolute coordinates.

Being sequencing-based, DNA microscopy is amenable to discovery of sequence variation; hence, for a given mRNA it might be possible to interrogate several features such as location, expression levels, genetic variation, splicing isoforms and modifications. Finally, DNA tags could be used for spatial analyses of other macromolecules, including DNA and proteins. As one potential application to explore, DNA microscopy might complement other methods for 3D chromosome conformation analysis.

Darren J. Burgess

ORIGINAL ARTICLE Weinstein, J. A., Regev, A. & Zhang, F. DNA microscopy: optics-free spatio-genetic imaging by a stand-alone chemical reaction. *Cell* **178**, 229–241 (2019)
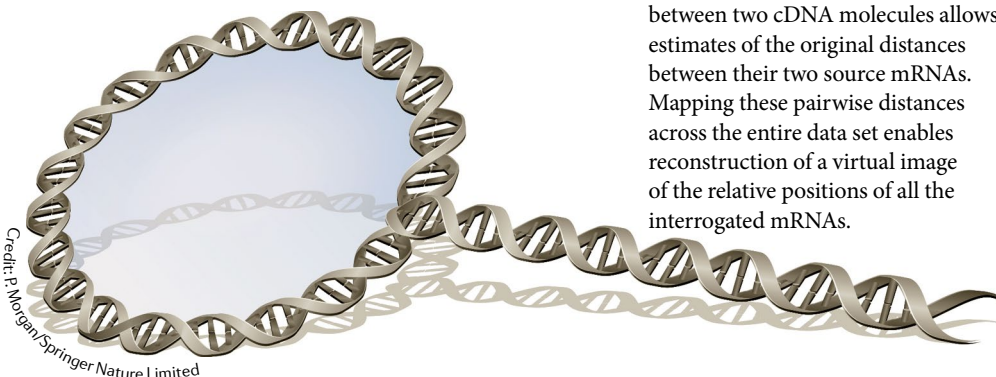
Credit: P. Morgan/Springer Nature Limited