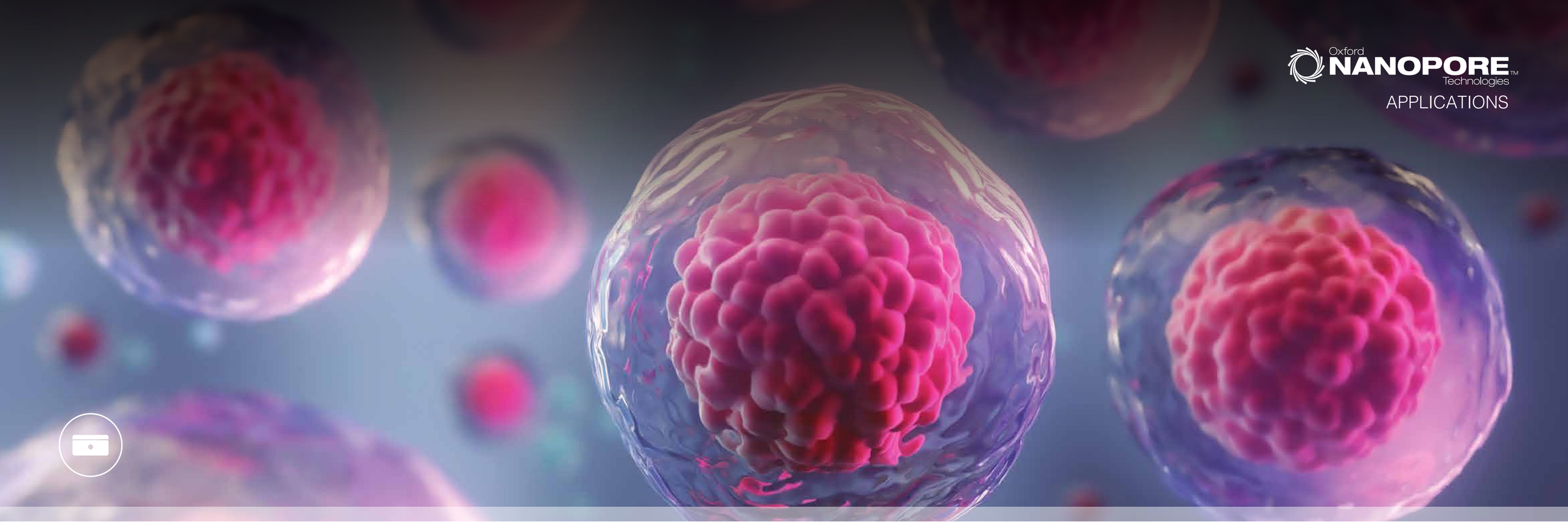


Oxford NANOPORE Technologies™

APPLICATIONS

Single-cell and spatial transcriptomics help to unlock our understanding of the subtleties of cellular diversity

A combination of single-cell and spatial platforms with full-length cDNA sequencing offers the potential to provide a level of detail to transcriptomic studies that is not available from bulk analyses

Contact: **apps_posters@nanoporetech.com** More information at: **www.nanoporetech.com** and **publications.nanoporetech.com**

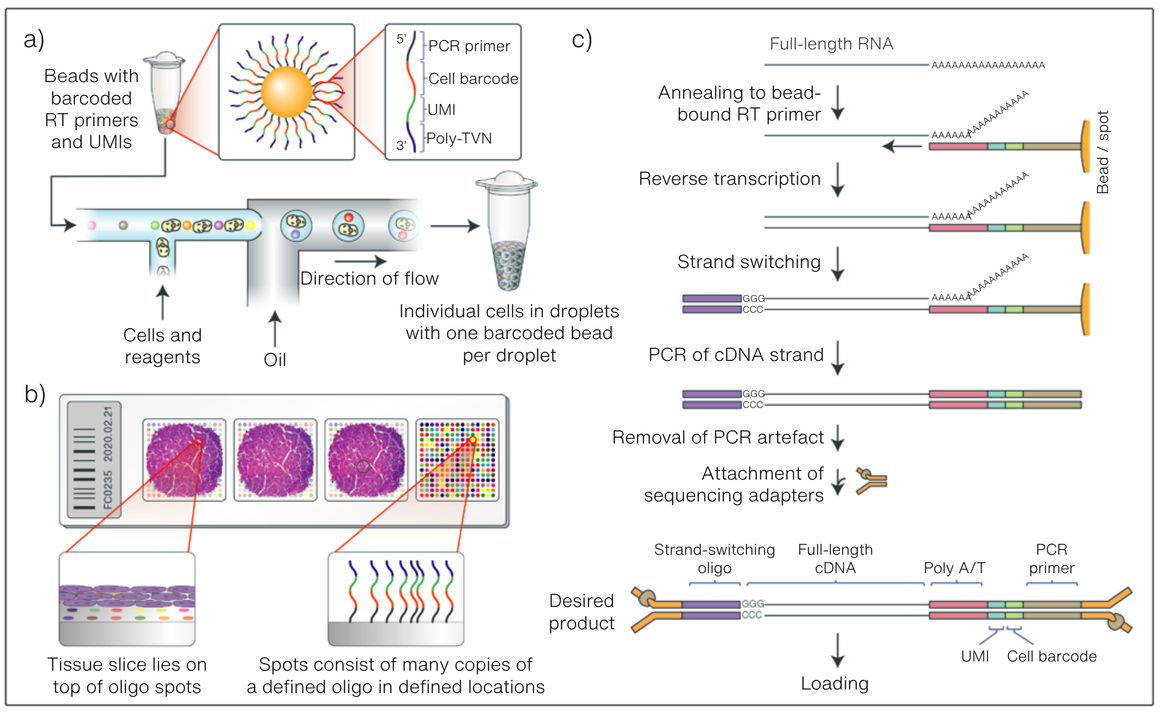
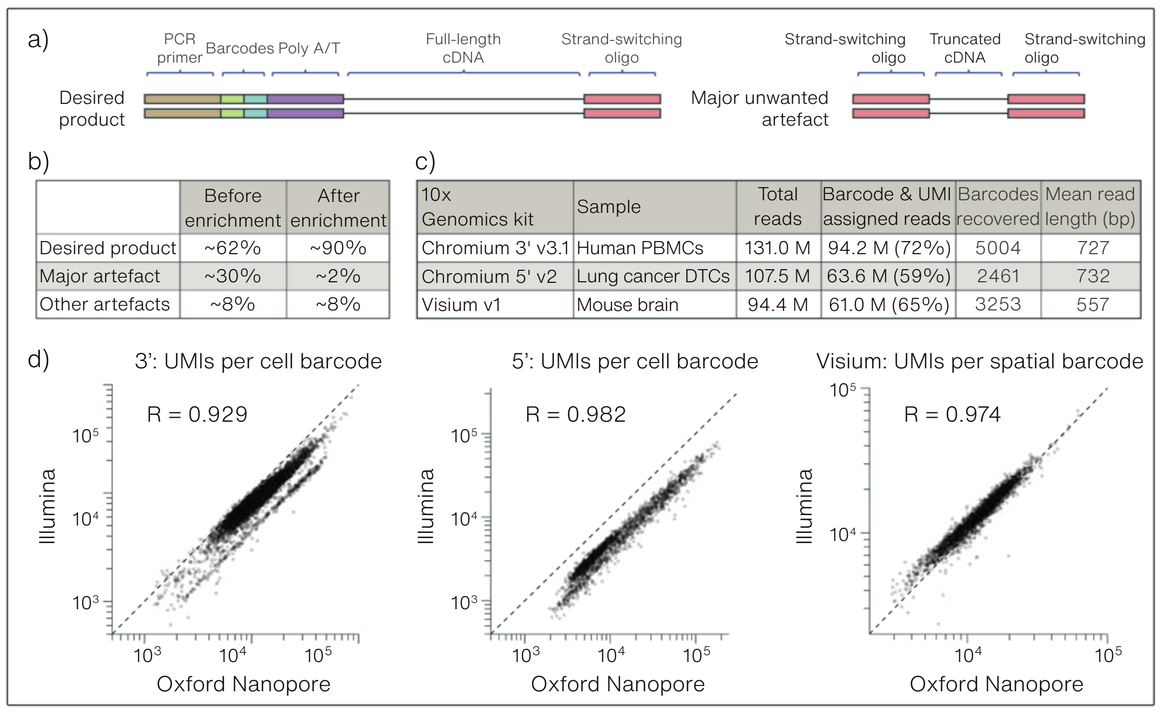


Fig. 1 Preparing a) single cell and b) spatial libraries for c) reverse transcription (RT)

Compatibility of existing single-cell and spatial platforms with nanopore sequencing

Differences in the transcriptomic behaviour of individual cells and across spatial dimensions of a sample are not visible in bulk analyses of heterogeneous cell populations. One way to compare expression levels of single cells is to encapsulate each cell in a droplet along with a bead coated with RT primers. The primers surrounding any one bead contain the same cell-barcode sequence. Cell lysis and RT occur within the droplets, so all cDNAs from the same cell are given the same cell barcode (Fig. 1a). For spatial transcriptomics, the concept is similar, with barcode-containing oligos bound to spots arrayed according to X & Y coordinates across a slide (Fig. 1b). Following RT and strand switching, cDNAs are pooled and amplified, and adapters are added (Fig. 1c).



	Before enrichment	After enrichment
Desired product	~62%	~90%
Major artefact	~30%	~2%
Other artefacts	~8%	~8%

10x Genomics kit	Sample	Total reads	Barcode & UMI assigned reads	Barcodes recovered	Mean read length (bp)
Chromium 3' v3.1	Human PBMCs	131.0 M	94.2 M (72%)	5004	727
Chromium 5' v2	Lung cancer DTCs	107.5 M	63.6 M (59%)	2461	732
Visium v1	Mouse brain	94.4 M	61.0 M (65%)	3253	557

Fig. 2 Enriching for full-length cDNAs in 10x Genomics libraries by biotin capture

Maximising full-length cDNAs in 10x Genomics single-cell and spatial libraries

PCR artefacts are frequently produced during amplification of barcoded single-cell cDNAs, limiting the proportion of full-length transcript reads. The 10x Genomics PCR artefact consists of a truncated cDNA flanked by copies of the strand-switching oligo (Fig. 2a). These can be depleted by biotin capture, giving a far higher proportion of full-length reads per run: ~90% (Fig. 2b). With this protocol, PromethION sequencing of libraries produced from various 10x Genomics kits generates >90 M reads, of which >60 M reads can be successfully assigned a barcode and UMI (Fig. 2c). We see high correlation between our single-cell expression levels and those from matching Illumina datasets across all supported 10x Genomics kits (Fig. 2d).

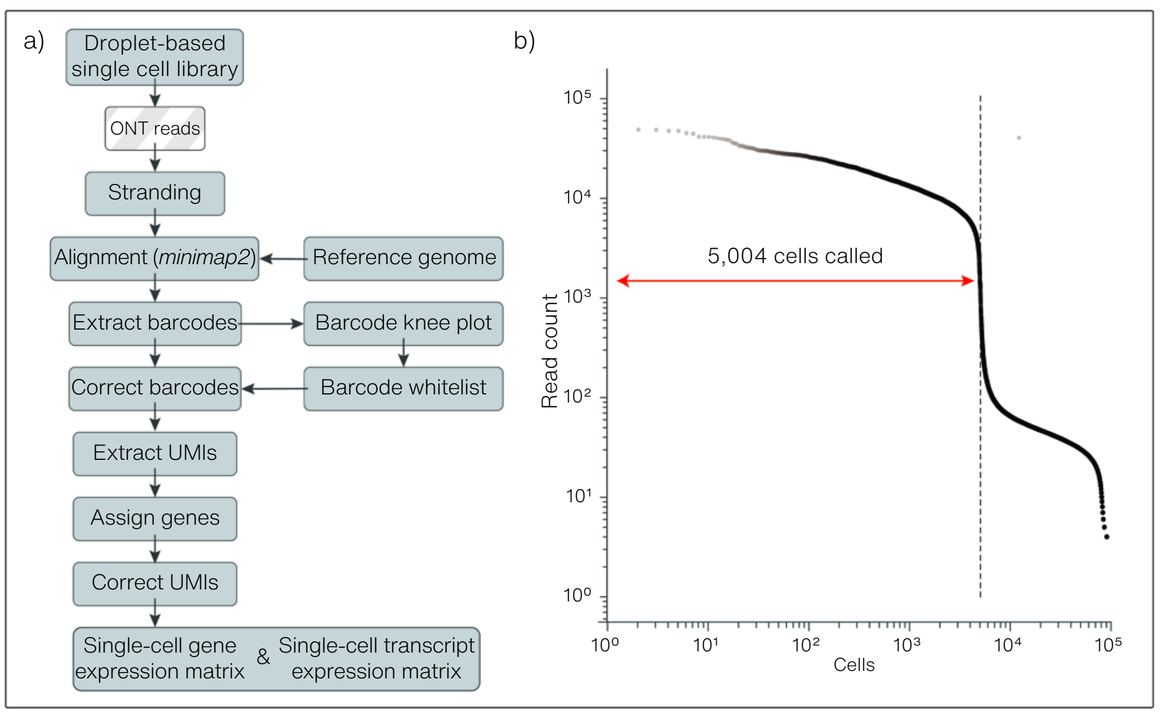


Fig. 3 Data analysis a) *wf-single-cell* workflow b) knee plot of a 3' human PBMC library

The *wf-single-cell* analysis pipeline enables ONT-only cell-barcode identification

Following data generation, putative cell barcodes from the highest quality reads are clustered using the *wf-single-cell* pipeline to identify the cell barcodes present in the sample (Fig. 3a). The resulting barcode clusters are visualised in a knee plot (Fig. 3b) to identify which cell barcodes have sufficient read support. Once these true barcodes are identified, cell barcodes are assigned to all reads based on edit-distance criteria. Barcode-assigned reads are then aligned and annotated using the appropriate references. The results can then be visualised through cell clustering based on gene expression, or full-length transcript consensus sequences can be generated to look at isoforms, alternative splicing and genotyping.

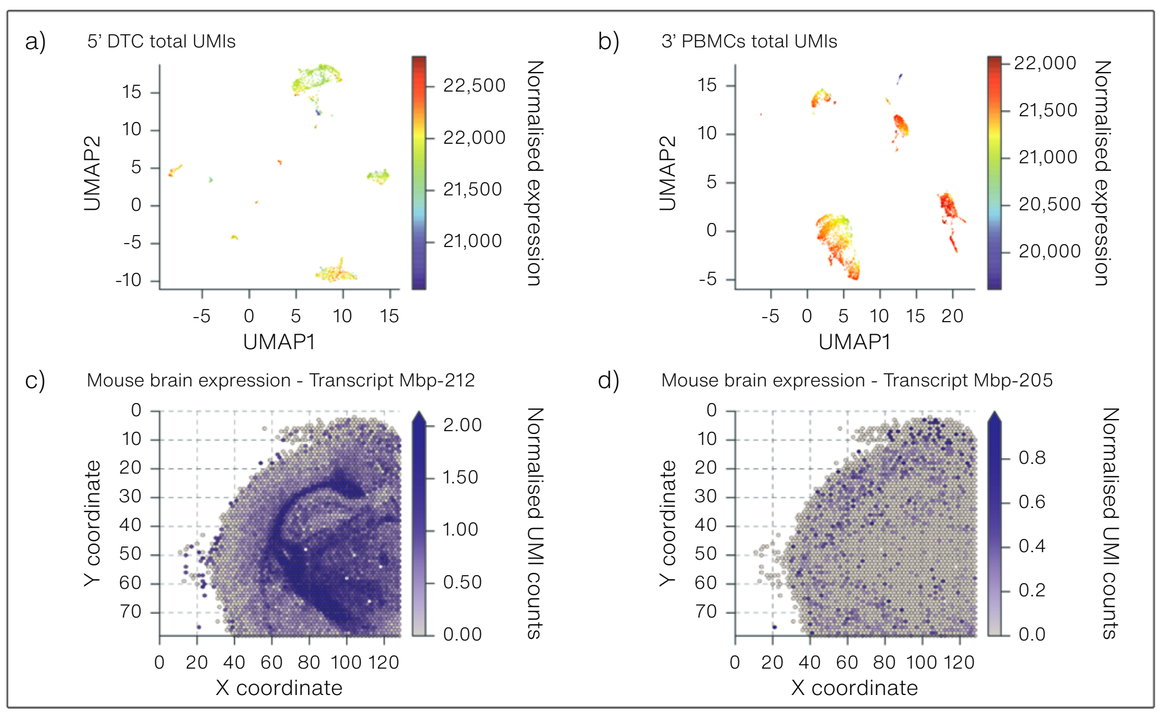


Fig. 4 UMI counts a) human DTCs and b) PBMCs c) mouse Mbp-212 and d) Mbp-205 isoforms

Gene and isoform expression from human single cells and a mouse brain spatial library

We generated UMAP plots from the single-cell gene expression of human lung carcinoma dissociated tumour cells (DTCs) using the 10x Genomics 5' single-cell gene-expression kit, revealing distinct cell types clustering together (Fig. 4a). We also profiled the gene expression of human peripheral blood mononuclear cells (PBMCs) using the 10x Genomics 3' single-cell gene expression kit, showing cells clustering by expression profiles (Fig. 4b). Finally, we used the spatial expression libraries produced by 10x Genomics' Visium to interrogate transcript-level spatial expression across a mouse brain sample. By sequencing full-length transcripts, we are able to uncover substantial differences in the expression of two isoforms from the Mbp gene, Mbp-212 (Fig. 4c) and Mbp-205 (Fig. 4d)