**1.10X genomics sequencing:** <https://www.10xgenomics.com/solutions/single-cell/>

An external file that holds a picture, illustration, etc.
Object name is ncomms14049-f1.jpg

Zheng G X, Terry J M, Belgrader P, et al. Massively parallel digital transcriptional profiling of single cells:[J]. Nature Communications, 2017, 8:14049.

(**a**) scRNA-seq workflow on GemCode technology platform. Cells were combined with reagents in one channel of a microfluidic chip, and gel beads from another channel to form GEMs. RT takes place inside each GEM, after which cDNAs are pooled for amplification and library construction in bulk.

(**b**) Gel beads loaded with primers and barcoded oligonucleotides are first mixed with cells and reagents, and subsequently mixed with oil-surfactant solution at a microfluidic junction. Single-cell GEMs are collected in the GEM outlet.

(**c**) Percentage of GEMs containing 0 gel bead (N=0), 1 gel bead (N=1) and >1 gel bead (N>1). Data include five independent runs from multiple chip and gel be ad lots over >70k GEMs for each run, n=5, mean±s.e.m.

(**d**) Gel beads contain barcoded oligonucleotides consisting of Illumina adapters, 10x barcodes, UMIs and oligo dTs, which prime RT of polyadenylated RNAs.

(**e**) Finished library molecules consist of Illumina adapters and sample indices, allowing pooling and sequencing of multiple libraries on a next-generation short read sequencer.

(**f**) CellRanger pipeline workflow. Gene-barcode matrix (highlighted in green) is an output of the pipeline.

主要概念流程：

**凝胶微珠**：英文叫“gel beads”。每个凝胶微珠上，种上特定的DNA片段。 每个DNA序列，分几段。

**Barcode**：这段Barcode一般是十几个碱基的长度。一个微珠是对应于一种Barcode，通过这400万种Barcode，可以把凝胶微珠给区分开。任意两个Barcode之间至少差两个或两个以上的碱基(两个hamming距离)，这样可以避免因为测序的时候对碱基的误读，而导致把两个Barcode搞混。

**UMI序列**：UMI是“unique multiplex index”的简称。UMI是一段随机序列，也就是说每一个DNA分子，都有自己的UMI序列。10个碱基长的UMI，有100万种序列的变化（4^10 = 1,048,576）。UMI起到的作用，是在经过PCR、深度测序得到的reads，可以看出哪些reads是来自于一 个原始cDNA分子的。这样，就可以把起始于一个原始cDNA分子，因为PCR扩增而产生的多个reads，简并成一个原始的cDNA分子。也就是可以排除各种cDNA，因为PCR扩增效率的 不同，而导致最后reads数量的偏差，也就是排除“PCR bias”。

**Poly(dT)序列**：这段序列起到的作用是与mRNA的Poly(A)尾巴结合，作为逆转录的引物，逆转录出cDNA来。

**芯片上的液流管路**：细胞混悬液在第一个十字交叉口，与凝胶微珠混合到一起，然后进入第二个十字交叉口，油相在这个十字交叉口加入进来。油把凝胶微珠和细胞的混悬液包裹 成一个又一个的油包水的小液滴，这些小液滴里面是水相，外面包裹的是油相。从总体上来说，这许多油包水的小微滴，就组成了一个乳浊液。在得到乳浊液之后，接下来把细胞膜破掉，让细胞当中的mRNA游离出来。游离出来的mRN A与小液滴中的水相混合，也就是和逆转录酶、结合在凝胶微珠上的核酸引物、以及dNTP底 物相接触。接着，发生逆转录反应。mRNA与凝胶微珠上带标签的DNA分子相结合，在逆转录酶的作用 下，逆转录出cDNA来。 注意，这样得到cDNA分子是带有这个微珠所特有的Barcode标签的，并且每个cDNA分子各 自还会带有特定的UMI标签。有了这个Barcode标签和UMI标签之后，这个cDNA分子就与其 它的cDNA分子区分开来。

**2.Seurat:** <http://satijalab.org/seurat/>

Seurat is an R package designed for QC, analysis, and exploration of single cell RNA-seq data. Seurat aims to enable users to identify and interpret sources of heterogeneity from single cell transcriptomic measurements, and to integrate diverse types of single cell data.

Input: Expression matrices

Main functions:

Data QC

Data filtration

Calculation of high-variance genes: choose high dispersion genes

Dimensional reduction: PCA

Graph-based clustering: K-nearest neighbor (KNN) graph

Identification of cluster markers: Choose diff-genes between cell clusters

**3.Moncle:** <http://cole-trapnell-lab.github.io/monocle-release/>

**Pseudotime**:

Build single-cell trajectories with the software that introduced pseudotime. Find cell fate decisions and the genes regulated as they're made.

**Clustering**:

Group and classify your cells based on gene expression. Identify new cell types and states and the genes that distinguish them.

**Differential expression**:

Find genes that vary between cell types and states, over trajectories, or in response to perturbations using statistically robust, flexible differential analysis.