The use of single-cell RNA-Seq to understand virus–host interactions

Sara Cristinelli and Angela Ciuffi

Single-cell analyses allow uncovering cellular heterogeneity, not only *per se*, but also in response to viral infection. Similarly, single cell transcriptome analyses (scRNA-Seq) can highlight specific signatures, identifying cell subsets with particular phenotypes, which are relevant in the understanding of virus–host interactions.

Address
Institute of Microbiology, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

Corresponding author: Ciuffi, Angela (angela.ciuffi@chuv.ch)

Introduction

Recent technological advances coupled with mathematical modeling and computer science revolutionized the field of biology. Indeed, deep sequencing of the whole human genome in 2001 signed a first milestone in the 'omics era' [1,2]. From that point on, lot of efforts have been devoted to develop new chemistries and devices with high throughput capacity, such as Next Generation Sequencing (NGS) and Mass Spectrometry (MS), and to bring them to basic and clinical research [3••,4]. The combination of multiple omics is known today as the 'integrome' and should allow a better and more comprehensive understanding of biological processes or diseases [5].

To date, a series of state of the art techniques is available for studies at population level resolution, with many biological applications, including cancer [6], stem cells [7], and infectious diseases [8,9]. These omic studies are likely to be valuable assets for personalized medicine [10–12].

Although essential and informative, population studies have some limitations as they contain a mixture of cells. As such, these cells can appear on one hand as being similar or homogeneous for some parameters, such as cell size or shape, tissue localization, protein expression or cell type subset. On the other hand, a cell population can also appear heterogeneous for some other parameters, such as cell cycle or metabolic activation. Thus, analyses at the cell population level may average and minimize individual cellular differences, potentially masking rare cells or cell subsets with a significant specific phenotype [13•]. Illustrations of this can be found in cancer, where heterogeneity in intra-tumor cells at genetic, epigenetic and phenotypic level can lead to resistance in cancer therapies, as well as in infectious diseases where cell heterogeneity can reveal differential susceptibility to infections or different immunological responses [14••,15,16]. In order to start investigating this cellular heterogeneity, single-cell analyses were initially performed using imaging techniques and further developed to benefit from sequencing technologies [14••,17]. Single-cell analyses bring up several technical challenges, including single cell isolation, handling minimal amounts of biological sample (6 pg of DNA, 10 pg of RNA), and data normalization and analysis [5,17–22]. Nowadays, an array of kits and methods are available, overcoming these challenges and allowing single cell omics analyses (reviewed in [23–26]).

Characterization of cell heterogeneity in the field of virology is obviously of great interest as it is likely impacting virus replication cycle and thus infection outcome. Indeed, the success of viral replication is highly dependent on its host, as viruses adapted to exploit the cell machinery to their own benefit. Interestingly, achieving universal infection in a given cell population is quite rare. Two non-mutually exclusive hypotheses can explain this unequal susceptibility to infection: (i) viral heterogeneity, that is, the presence of a mixture of competent, mutated and defective viral particles, displaying differences in infection ability, and (ii) cellular heterogeneity, that is, the presence of a mixture of cells with differences in metabolism, composition, activation status or cell cycle, resulting in specific cellular environments shaping virus progression success throughout the cell. In this perspective, single-cell analyses represent novel opportunities to identify specific cellular and molecular features promoting or in contrast restricting virus replication, thereby adding to the understanding of virus–host interactions, as well as providing new targets to inhibit viral replication.

To date, only a limited amount of single cell studies in virology have been published (reviewed in [14••,27]).

Initial studies used single-cell technologies to explore virus evolution, through methods including fluorescent

reporters, FACS analysis and sorting, as well as time lapse microscopy (recently reviewed in [27]), and it is only recently that transcriptome analyses at single cell level started to uncover cell subsets with specific response to viral infection. In this review, we will focus on the prominent role of single-cell RNA-Seq (scRNA-Seq) in virus–host interactions. In particular, we will provide an overview of the complete scRNA-Seq workflow as well as studies using this technology to identify specific features relevant for virus–host interactions.

Transcriptome analysis: single cell RNA-Seq workflow

The scRNA-Seq analysis pipeline can be divided in three essential steps (Figure 1): (i) single cell isolation, (ii) RNA-Seq procedure, that is, RNA extraction, reverse transcription (RT), library preparation and sequencing, and (iii) data analysis.

Single cell isolation

In order to perform single-cell RNA-Seq (scRNA-Seq) the first trivial step is to isolate single cells. To date single cell isolation methods can be divided according to two criteria: (i) the cell isolation capacity, that is, high or low throughput, depending of the level of automation and (ii) the cell selection method, that is, blind selection without any *a priori* or biased selection according to a specific parameter (Table 1). For example, some high throughput technologies such as Fluorescence Activated Cell Sorting (FACS) and Magnetic Activated Cell Sorting (MACS) allow a biased selection of the cell population based either on cell size and shape or on the expression of specific cell surface markers, while microfluidics-based and microdroplets-based technologies allow unbiased separation of single cells. Similarly, low throughput techniques that require micromanipulation, such as manual cell picking or Laser Capture Microdissection (LCM), also allow biased cell selection based on a specific cell morphology or the expression of a fluorescence marker. By contrast, limiting dilution-based methods are unbiased. (For more information on the isolation methods, please refer to [27–29,30••].)

RNA-Seq procedure: RNA extraction, reverse transcription, library preparation, and sequencing

Standard protocols for library preparation have been optimized to use only a few ng (10–100 ng) of starting DNA material. However, one single cell contains on average only ∼10 pg of total RNA. Therefore, workflows for RNA extraction and library preparation were adapted and optimized to work with single-cell material.

Isolated single cells need first to be lysed in order to access RNA. This step can be performed through dedicated automated devices (such as fluidigm C1, 10X genomics chromium, ddSEQ single-cell isolator), or manually, using commercially available kits (such as GenElute^{TM}

Single Cell RNA Purification Kit (Sigma) or Single Cell Lysis Kit (Thermo Scientific)). Alternatively, manual cell lysis and total RNA purification can be performed simultaneously using resin-based columns (e.g. Single Cell RNA Purification Kit (NorGen Biotek Corp), PicoPure® RNA Isolation Kit (Thermo Scientific)) or magnetic bead separation (e.g. 5 min Single Cell RNA Extraction Kit (Biofactories)). Afterwards, RNA is enriched for mRNA, either by polyA selection or more rarely by ribosome depletion (Table 2, Figure 1) [31,32]. The enriched RNA fraction is then usually reverse transcribed with modified oligodT primer through different protocols, either polyA tailing or template switching. During the RT step, some protocols allow tagging of single molecules with unique molecular identifiers (UMI), which are random hexanucleotides that can be used to quantify more precisely the number of initial mRNA molecules present in one single cell [33,34]. Following reverse transcription, cDNA is amplified by *in vitro* transcription or by PCR. The amplified cDNA library is then used for library preparation and high throughput sequencing (currently most protocols use Nextera for library preparation and Illumina as a sequencing platform) (for additional information on RNA-Seq procedures, please refer to [3••,15,23,25,29,35,36••,37•]).

Single-cell data analysis

As every single cell is unique, it is not possible to perform experimental replicates and assess noise. It is thus essential to have some quality controls to ensure data reliability. This can be achieved by adding synthetic mRNAs of known sequence and amount, such as External RNA Controls Consortium (ERCC) RNA spike-ins, to each cell lysate. The number of reads recovered from the spike-ins will provide information about inter-sample technical variability [3••,25,38•,39].

It has been recently shown that sequencing errors in UMI sequence are common [34], and may thus bias transcript quantification when used. To minimize the error rate of UMI during sequencing, Lau *et al*. developed a new approach consisting in building a high error resistant exponentially-expanded barcodes (EXBs) with giga-scale diversity and introduce them into double strand cDNA by mean of a transposase [40]. Multiple packages for single-cell data analysis are already available and many computational tools are being developed to improve current analyses in order to handle *bona fide* cell-to-cell variation and technical noise [38•,39,41,42,43••,44,45••,46]. Most of these analyses rely on cell-to-cell vicinity, or transcriptome similarity, thereby allowing a neighborhood-based graphical representation of individual cells, such as Principal Component Analysis (PCA) or t-distributed Stochastic Neighbor Embedding (t-SNE) plots. These plots are commonly used as they allow easy and intuitive visualization of cell heterogeneity and cell clusters that can be further analyzed by differential expression analysis.

Figure 1

Current Opinion in Virology
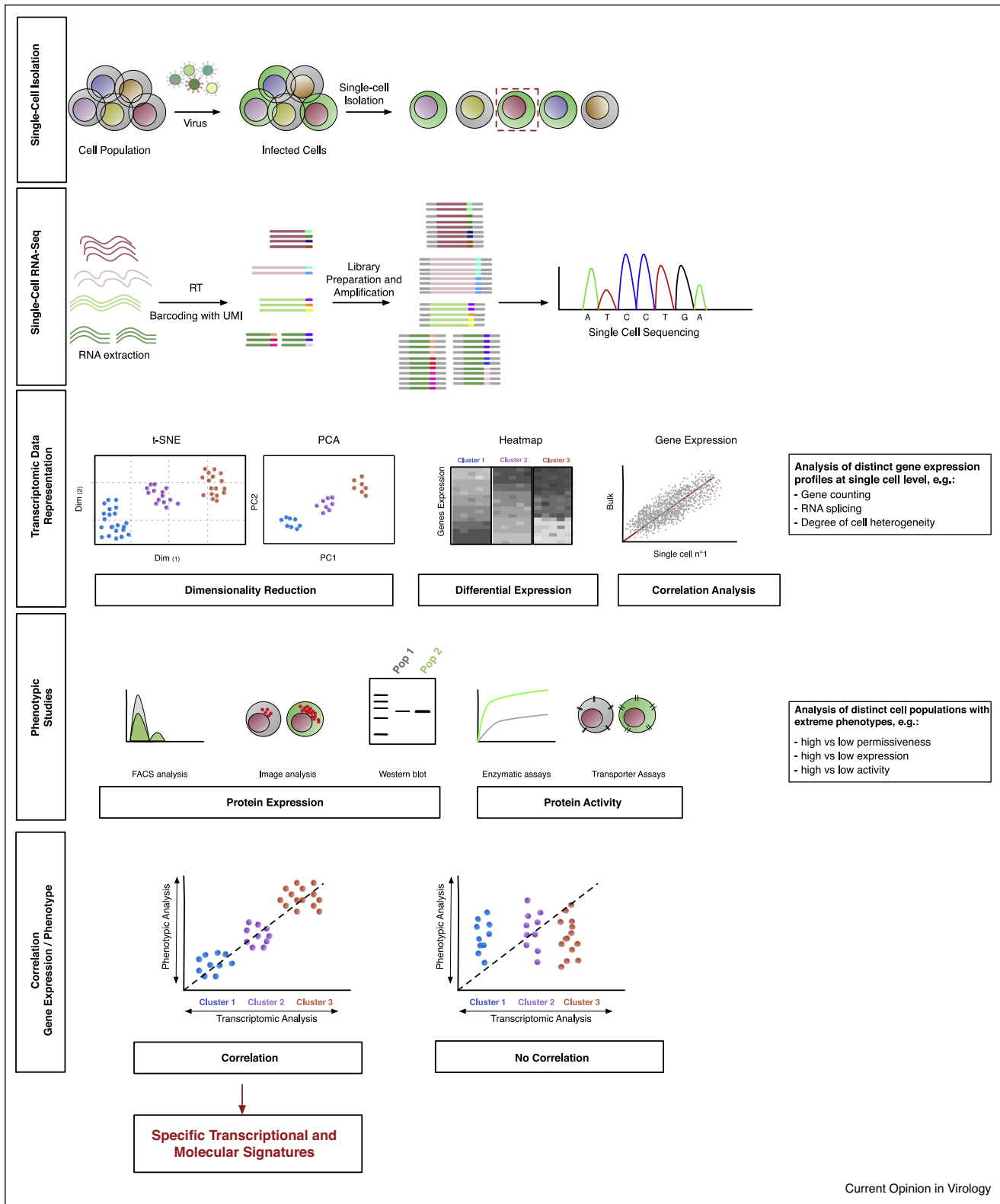
Identification of specific cell signatures using single-cell RNA-seq and phenotypic analyses. The whole analysis pipeline can be divided in five steps. (i) Cells infected at population level are separated and isolated as single cells. (ii) Each individual cell is processed for RNA-Seq, requiring RNA extraction, reverse transcription (RT) with optional unique molecular identifier (UMI) barcoding, library preparation and amplification, and

Table 1

Single cell isolation method.

Throughput	Technique	Cell selection	Parameters of population enrichment	Number of cells	Final volume/ cell
High	FACS	Biased	Cell morphology/size, specific cell surface marker expression, reporter gene expression	Hundreds	Microliters
	MACS	Biased	Specific cell surface marker expression	Hundreds	Microliters
	Microfluidics (i.e. Fluidigm C1, 10X genomics chromium)	Unbiased	–	Hundreds to thousands	Nanoliters
	Microdroplets (i.e. ddSEQ Single-Cell Isolator)	Unbiased	–	Thousands	Nanoliters
Low	Micromanipulation (manual cell picking, laser capture microdissection)	Biased	Cell morphology reporter gene expression	Tens	Microliters
	Limiting dilution	Unbiased	–	Tens	Microliters

In conclusion, transcriptome analysis at single-cell level (scRNA-Seq) can reveal cell heterogeneity in a cell population, thereby highlighting the existence of one or multiple discrete cell clusters with specific gene expression profiles.

Identification of a cell-specific transcriptional signature correlating with a specific phenotype in response to viral infection

Although valuable, scRNA-Seq alone does not provide any information on the association and relevance of a specific cell cluster (with a specific gene expression signature) with a specific cell phenotype. For this purpose, identification and assessment of a specific phenotype in response to viral infection is required.

Phenotypic analyses usually rely, either on the expression of a virus reporter gene or a specific cellular gene, or on the activity of a specific cellular protein (Figure 1). Furthermore, phenotypic analyses can be performed either at single-cell level associated with scRNA-Seq or at population level independently of scRNA-Seq. Indeed, assessment of the phenotype of interest on the same single cell that will be lysed and processed for scRNA-Seq allows for direct overlapping of transcriptome and phenotype, thereby providing a direct and straightforward association between gene expression signature and specific phenotype. This is possible if single cells are first characterized by FACS or by microscopy for the expression of a fluorescent reporter gene before scRNA-Seq for example. Alternatively, specific phenotypes of interest need to be first characterized at population level. In this case, it will be helpful to identify two cell populations with extreme opposite phenotypes. For example, a cell population A is more permissive than a cell population B. Single-cell analysis of the two cell populations may then reveal imbalanced cell clustering, correlating with the observed phenotypes. For example, the cell population A contains a higher proportion of cells in cluster 3 and only few cells in the other two clusters, while cell population B contains only a few cells in cluster 3 and a higher proportion of cells in the other clusters. Cells in cluster 3 will thus represent permissive cells and will be characterized by a specific transcriptional signature.

In conclusion, joint analysis of scRNA-Seq profiles and phenotype assessment brings additional and valuable information, as it allows correlating a specific cell phenotype to a transcriptional signature (Figure 1). These analyses should help identifying novel biomarkers specific for viral phenotypes.

(Figure 1 Legend Continued) finally sequencing. (iii) Sequencing reads are further processed and aligned to host and virus genome reference sequences to assess abundance and structure of transcripts, thereby defining the gene expression profile of each single cell. Multiple transcriptomic analyses can be performed to inform about data structure and cell heterogeneity using dimensionality reduction plots, such as t-Distributed Stochastic Neighbor Embedding (t-SNE) or Principal Component Analysis (PCA). Cell clustering and differential expression analysis can provide specific gene expression profiles. Finally, single cell transcriptomes can be compared to cell population transcriptome or to quantitative RT-qPCR analyses. These three first steps characterize the process for single-cell transcriptomic analysis, allowing mostly revealing the level of cellular heterogeneity in a cell population, that is, identifying one or multiple cell subpopulations. (iv) Phenotypic analyses at single cell level (with direct association with corresponding scRNA-Seq) or using two cell populations with extreme opposite phenotypes can complement transcriptome analyses. Phenotypic analyses include assessment of protein expression level and localization, using FACS staining, immunofluorescence or western blots, or assessment of protein activity studies, using assays quantifying enzymatic activity, transport activity (channels) or translocation assays (metabolites). The analysis of specific phenotypes should also reveal single cell heterogeneity (i.e. single cells displaying different levels of protein expression or activity) or differences between multiple cell populations. (v) Finally, correlation analysis between transcriptomic data and phenotypical analyses should help identifying specific cell subsets displaying a specific phenotype and characterized by a defined transcriptional and molecular signature.

Table 2

ScRNA-Seq workflow.

RNA-Seq procedure[a]	Advised single cell isolation method	Starting material	RT	UMI[b]	Amplification	cDNA coverage
mRNA-Seq (Tang's method) [61••]	Any	PolyA selected RNA	PolyA Tailing	No	PCR	Full-length with 3′-end biases
Quartz-Seq [62]	Any	PolyA selected RNA	PolyA Tailing	No	PCR	Full-length with 3′-end biases
SC3-Seq [63]	Any	PolyA selected RNA	PolyA Tailing	No	PCR	Full-length with 3′-end biases
SUPer-Seq [31]	Any	Ribosome depleted RNA	Poly-dA Tailing	No	PCR	Full-length
MATq-Seq [32]	Any	Ribosome depleted RNA	Poly-dC Tailing	Yes	PCR	Full-length
SMART-Seq [64]	Any	PolyA selected RNA	Template Switching	No	PCR	Full-length
SMART-Seq 2 [65]	Any	PolyA selected RNA	Template Switching	No	PCR	High coverage full-length
STRT-Seq [66]	Microdroplet	PolyA selected RNA	Template Switching	Yes	PCR	Fragments <500 bp 5′-end (TSS) biases
Drop-Seq [67]	Microdroplet	PolyA selected RNA	Template Switching	No	PCR	Strong 3′ end biases
Cel-Seq [68]	Microdroplet	PolyA selected RNA	PolyA Tailing	Yes	*In vitro* transcription (IVT)-based	Strong 3′ end biases
Cel-Seq 2 [69]	Microdroplet	PolyA selected RNA	PolyA Tailing	Yes	*In vitro* transcription (IVT)-based	Strong 3′ end biases
In-Drop [70]	Microdroplet	PolyA selected RNA	PolyA Tailing	Yes	*In vitro* transcription (IVT)-based	Strong 3′ end biases
MARS-Seq [71]	Microdroplet	PolyA selected RNA	PolyA Tailing	Yes	*In vitro* transcription (IVT)-based	Strong 3′ end biases

[a] SC3-Seq (single-cell mRNA 3-prime end sequencing); MATq-Seq (multiple annealing and dC-tailing-based quantitative single-cell RNA-Seq); SMART-Seq (Switching mechanism at 5' End of RNA Template); STRT-Seq (Single cell tagged reverse transcription sequencing); Cel-Seq (Cell Expression by Linear amplification); MARS-Seq (Massive parallel RNA single cell sequencing framework).
[b] Unique Molecular Identifiers.

Single-cell RNA-Seq in virology

To date, a few single-cell RNA-Seq studies have been performed, investigating virus or cell diversity (Table 3). In 2013, McWilliam Leitch and McLauchlan published the first scRNA-Seq analysis in virology, aiming at exploring the heterogeneity of Hepatitis C virus (HCV) quasispecies and assessing their fitness [47]. They isolated 16 Huh7 single cells that express constitutively JFH1_2a subgenomic replicon, extracted total RNA and amplified a 708 bp-specific N5B region by RT-qPCR. The PCR products were then cloned and 20 clones per single cell were sequenced, providing a total of 320 viral sequences. They showed that one single cell harbored on average 113 copies of viral RNA and that the sequence of HCV quasispecies differed extensively among cells, with some single cells displaying only WT viral RNA sequences while others contained up to four different viral RNA sequences, and diverse mutations. These data highlighted the co-existence of multiple quasi-species within the same single cell, suggesting distinct individual viral evolution, likely due to mutations introduced during genome replication by the error-prone RNA-dependent RNA polymerase. Furthermore, fitness analysis of three major variants was compared to the prevalent WT replicon and showed that the best replicative fitness was displayed by the WT replicon, consistent with a selective advantage for WT replication and its high prevalence.

In 2015 Wu *et al*. developed a dedicated pipeline called mRNA amplification and library construction system (MIRALCS) to investigate both cellular and viral RNA of the tumoral HeLa S3 cell line, originally transformed by Human Papilloma Virus (HPV) [48]. scRNA-Seq analysis revealed a high degree of heterogeneity among individual virus-infected cells. In particular, they identified alternative splicing of HPV genome, resulting either in internal E6 splicing leading to a truncated and non-functional E6 protein, or in aberrant HPV-host sequence fusion transcripts. Furthermore, coupling transcriptomic data with co-expression analysis identified a set of 283 genes co-regulated with E6 and E7, representing novel potential HPV interactors.

Tsioris *et al*. used microengraving to identify and capture West Nile Virus (WNV)-specific actively secreting and memory B cells isolated from 11 WNV-infected individuals, followed by single cell RT-PCR and sequencing of WNV-specific antibodies (V, D, J, CDR3 sequences of heavy and light immunoglobulin chains) [49,50]. They successfully identified and isolated 98 WNV-specific cells out of thousands of cells and identified complete antibody sequence from 19 single cells. Phenotypic analyses were carried out to characterize the diverse antibodies, using Plaque Reduction Neutralization Assay and binding affinity to WNV Envelope protein. This strategy successfully led to the identification of four novel WNV neutralizing antibodies.

Afik *et al*. developed T cell receptor (TCR) Reconstruction Algorithm for Paired-End Single cell (TRAPeS), an algorithm facilitating reconstruction of full-length TCR sequences from scRNA-Seq-derived short reads, in order to investigate TCR heterogeneity among CD8+ T cells [51]. Their analysis revealed that naïve CD8+ T cells expressed distinct TCR, thereby suggesting a heterogeneous population, while effector memory CD8+ T cells displayed some levels of oligoclonality. Similarly, virus-exposed CD8+ T cells, during vaccination protocols or chronic infections, presented various degrees of heterogeneity. Indeed, Yellow Fever Virus (YFV) and Hepatitis C virus (HCV)-exposed CD8+ T cells (short-term exposure during vaccination protocol) displayed higher levels of TCR heterogeneity than cytomegalovirus (CMV)-exposed cells (long-term exposure during persistent infection). These data suggest that TCR heterogeneity and cell oligoclonality were negatively correlated and depended on the duration of virus exposure. ScRNA-Seq data of 353 single cells were further explored to assess cell heterogeneity of CMV and YFV-exposed CD8+ T cells. The analysis clearly separated clusters of naïve CD8+ T cells, effector CD8+ T cells and CMV-specific CD8+ T cells. By contrast, YFV-specific CD8+ T cells displayed some cellular heterogeneity, with one effector memory-like subpopulation and one naïve-like subpopulation. Further analysis revealed that YFV-specific CD8+ T cell subpopulations were associated with a distinct TCR sequence, displaying longer CDR3 sequences in the naïve-like subset.

Recently, two studies used scRNA-Seq to investigate Zika virus (ZIKV)-mediated neuropathogenesis [52,53]. First, Nowakosky and colleagues used scRNA-Seq to characterize the gene expression profile of multiple cell types of the developing human cortex, in order to correlate ZIKV tropism and pathogenesis [52]. In particular, they evaluated the expression of 23 gene candidates potentially involved in ZKV across cell types and identified AXL as being highly expressed in radial glial cells, in astrocytes, in microglia and in endothelial cells. Further analysis of AXL expression was performed by immunostaining of multiple tissues, including iPSC-derived cerebral organoids, highlighting AXL expression in radial glia (ventricular zone staining) and neural stem cells. AXL expression pattern seemed thus consistent with cells susceptible to ZIKV infection and cerebral lesions (microcephaly and ocular lesions), further arguing for a possible role of AXL in ZIKV entry.

Onorati *et al*. established cells lines derived from neuroepithelial stem (NES) cells isolated from primary neocortex (NCX) or from spinal cord (SC) in order to investigate ZIKV infection success according to tropism and gene expression profile (scRNA-Seq) [53]. They confirmed previous data showing high expression of AXL in NES cells and radial glial cells, as well as preferential localization of ZIKV-infected cells at the ventricular zone. ZIKV infection of

Table 3

ScRNA-Seq studies in virology.

Virus[a]	Cell type	Single cell isolation	Nb of single cells analyzed	RNA sample	RNA-Seq method	Additional transcriptomic analysis	Phenotypic analysis	Reference
HCV	Huh-7 cell line, expressing JFH-1_2a strain	Limiting dilution	16	vRNA	RNA extraction: RNeasy micro kit (Qiagen); gene-specific qRT-PCR (N5B); Cloning into sequencing vectors	Bulk RNA-Seq; RT qPCR	Replicative fitness analysis	McWilliam Leitch and McLauchlan, 2013 [47]
	Key findings/outcome: – Presence of an average of 113 copies of HCV replicons per cell – Assessment of viral sequence heterogeneity (i.e. viral quasispecies) in each single cell, present at various frequencies – Distinct single cells provide local compartments with different selective environments shaping distinct viral sequence evolution – The predominant quasispecies in each single cell was the one displaying the best fitness							
HPV	HeLa S3 cell line, transformed by HPV infection	MIRALCS	40	Cell RNA and vRNA	MIRALCS; SMART-Seq2 (modified)	Bulk RNA-Seq RT-qPCR; Transcriptomic co-expression analysis	Cell cycle	Wu et al., 2015 [48]
	Key findings/outcome: – Average number of mRNA molecules in HeLa S3 was 152 000 – Cell heterogeneity at cell cycle (G2/M and non-G2/M cells) – Cell heterogeneity at the transcriptome level, with differences in gene expression, alternative splicing isoforms, and fusion transcripts – Splice variant heterogeneity in HPV sequences, resulting in truncated E6 proteins or in HPV-host chimeric transcripts – Identification of 283 E6–E7 co-regulated genes as novel candidates potentially interacting with HPV proteins							
WNV	Primary memory B cells and antibody-secreting cells, isolated from infected individuals	Micro-engraving	19	Cell RNA (antibody coding sequence: V, D,J; CDR3)	Micro-engraving; gene-specific single-cell RT-PCR	Bulk RNA-Seq	Immunoassay to identify WNV-specific cells WNV Ab characterization (Plaque reduction neutralization assay, Immunoblot and antigen affinity assays). Humoral response analysis	Tsioris et al., 2015 [49]
	Key findings/outcome: – Identification of four novel WNV neutralizing antibodies, with determined coding sequence – Frequency assessment of WNV-specific B cells (∼30 cells per 100 000 peripheral blood mononuclear cells) – Persistence of WNV antibodies and no predictive value for disease severity							

Table 3 (*Continued*)

Virus[a]	Cell type	Single cell isolation	Nb of single cells analyzed	RNA sample	RNA-Seq method	Additional transcriptomic analysis	Phenotypic analysis	Reference
YFV CMV HCV	Human and murine CD8+ T cells	FACS sorting	513	Cell RNA (TCR repertoire)	SMART-Seq2	Bulk RNA-Seq	Single cell Surface marker expression (FACS); Tetramer binding assay (fluorescence staining)	Afik et al., 2017 [51]

Key findings/outcome:
– Analysis of TCR repertoire diversity
– Naïve CD8+ T cells express unique TCR, YFV- and HCV-specific CD8+T cells displayed a higher TCR repertoire diversity compared to CMV-specific CD8+ T cells which are oligoclonal (poor TCR diversity)
– Correlation between TCR diversity, oligoclonality, and duration of virus exposure
– Identification of two YFV-specific CD8+ T cell subsets (effector memory-like and naïve-like)
– Correlation between YFV-specific CD8 +T cells differentiation state and CDR3 length (naïve cells have longer CDR3 sequences)

| ZIKV | Primary cells from human cortex (astrocytes, radial glia, intermediate progenitor cells, neurons), iPSC-derived cerebral organoids | Micro-dissection and Fluidigm | 715 | Cell RNA | SMARTer® Ultra™ Low RNA Kit | | Immunostaining of candidate markers (immunohisto-chemistry) | Nowakoski et al., 2015 [52] |

Key findings/outcome:
– Expression of AXL, a candidate gene involved in ZIKV entry, in radial glia, astrocytes, endothelial cells and microglia
– Characterization of AXL expression in cerebral and organoid samples

| ZIKV | Neuroepithelial stem (NES) cells Neurocortical cells and Radial Glia Cells | Fluidigm | 1009 | Cell RNA | SMARTer® Ultra™ Low RNA Kit | | Immuno-fluorescence staining; Immunoblot; Immunoelectron microscopy | Onorati et al., 2016 [53] |

Key findings/outcome:
– Establishment and scRNA-Seq characterization of neuroepithelial stem (NES) cell lines as model for *in vitro* ZIKV studies
– Expression and localization analysis of AXL and TYRO3, genes potentially involved in ZIKV entry in NES cells and Radial Glia
– ZIKV may impair mitosis and survival of NES cells by inducing the relocation of pTBK1 from mitotic centrosomes to mitochondria

| CMV | Glioblastoma tumors and glioblastomacell lines | Sequences downloaded from previous study ([55], SMART-Seq) | 854 | Cell RNA and vRNA | | Bulk RNA-Seq | | Johnson et al., 2017 [54] |

Key findings/outcome:
– Lack of CMV expression in glioblastoma tumors and cell lines

Table 3 (*Continued*)

Virus[a]	Cell type	Single cell isolation	Nb of single cells analyzed	RNA sample	RNA-Seq method	Additional transcriptomic analysis	Phenotypic analysis	Reference
HIV	Dendritic Cells (DCs) from HIV-1 Elite Controller (EC) individuals	FACS sorting	139	Cell RNA			FACS analysis on surface marker; Functional characterization of DCs; Validation on healthy cDCs	Martin-Gayo *et al.*, 2016 [56]
	Key findings/outcome: – Presence of 3 clusters in cDCs from Elite controllers – Characterization of a highly functional subset of cDCs with antiviral profile (high expression of Interferon-stimulated genes, costimulatory molecules and cytokines – High expression of CD64 and PD-L1, correlating with improved immune response to HIV								
HIV	Primary CD4+T cells, with extreme phenotypes of HIV permissiveness (high vs low permissive cells)	Fluidigm	166	Cell RNA	SMART-Seq2	Bulk RNA-Seq	Immunoprofiling; Cell surface protein expression (FACS immunoprofiling); Activation assays; Infection assays	Rato *et al.*, 2017 [57]
	Key findings/outcome: – CD4+ T cell heterogeneity was driven by T cell activation – Cell heterogeneity appeared as a continuum of transcriptional cell states rather than distinct clusters – Correlation analysis of scRNA-Seq, protein expression, activation and permissiveness to HIV identified markers of HIV permissiveness – HIV permissive cells displayed low levels of genes involved in interferon I pathway or in defense response pathway								
HIV	Primary model of latently infected CD4 + T cells	Fluidigm	224	Cell RNA and vRNA	SMART-Seq2	Bulk RNA-Seq	Fluorescence based protein expression analysis (FACS and microscopy)	Golumbeanu *et al.*, 2017 [58]
	Key findings/outcome: – Latent CD4+ T cells were heterogeneous, with two distinct cell clusters – Latent CD4+ T cells displayed differential response to TCR-mediated stimulation – Transcriptional heterogeneity was linked to viral reactivation potential								

[a] HCV: Hepatitis C Virus; HPV: Human Papilloma Virus; WNV: West Nile Virus; ZKV: Zika Virus; CMV: Cytomegalovirus; YFV: Yellow Fever Virus; HIV: Human Immunodeficiency Virus.

organotypic brain slices confirmed the absence of infected cells in the cortical plate and the preferential infection of cells residing in the ventricular zone such as radial glial cells. Over time, a progressive disruption of radial glia scaffold and disorganization of the neocortical structure was observed, consistent with brain damage. They also investigated the role of TNAK-binding kinase 1 (TBK1), previously involved in innate immunity and cell proliferation, in NES cell growth arrest and cell death. Although scRNA-Seq data did not reveal any difference in TBK1 gene expression, fetal cortex analyses revealed a higher level of phosphorylated TBK1 (pTBK1) in the ventricular zone. Furthermore, ZIKV infection relocated pTBK1 from mitotic centrosomes to the mitochondria in NES-infected cells, resulting in mitosis disruption and increased cell death, and consistent with ZIKV-mediated cerebral lesions.

Johnson *et al.* performed scRNA-Seq on glioblastoma tumors and glioblastoma cell lines, previously reported to be heterogenous, in order to investigate a possible role of CMV infection in glioblastoma [54,55]. The analysis of 854 single cells failed to detect CMV transcripts, suggesting either the absence or the rarity of CMV sequences. Therefore, these data indicated that glioblastoma tumors may not contain active CMV.

Recent studies on the Human Immunodeficiency Virus (HIV) used single-cell RNA-Seq to identify the transcriptional signature of specific cell subsets, such as dendritic cells from elite controllers, highly permissive CD4+ T cells or inducible latently infected CD4+ T cells [56–58]. Martin-Gayo *et al.* characterized conventional dendritic cells (cDCs) isolated from elite controllers, at the single-cell level, in order to understand if these cells contributed to the improved response observed in these individuals [56,59]. Using scRNA-Seq, they identified three cDC subsets after HIV exposure: one subset similar to control unexposed cells, one subset of exposed cDCs, and a third exposed cDC subset with a different expression profile, showing upregulated antiviral response (interferon-stimulated genes and cytokines) and effective antigen-presenting properties. This third functional subset was enriched in elite controllers compared to chronic progressors and uninfected individuals, and may thus contribute to the improved immunological control observed in elite controllers.

Rato *et al.* identified two CD4+ T cell populations (isolated from two individuals) that, upon TCR-mediated stimulation, displayed high and low permissiveness to HIV infection [57]. scRNA-Seq analysis of non-infected activated CD4+ T cells revealed transcriptional heterogeneity, with individual cells reflecting a continuum of cell states and correlating with cellular activation status. High permissive cells contained more individual cells with successful activation profile than low permissive cells. Analysis of cell surface protein expression was further assessed by FACS and correlated with

permissiveness to HIV. Correlation analysis of scRNA-Seq, cell activation, protein expression and HIV permissiveness identified candidate cellular markers marking HIV permissive cells. A few of these biomarkers were further validated as biomarkers of HIV permissiveness. Subsequent transcriptome analysis of biomarker-sorted cell subpopulations identified specific transcriptional features associated with supported viral replication. In particular, highly permissive cells displayed downregulated expression of 95 genes involved in antiviral response.

Using a similar approach, Golumbeanu *et al.* investigated the cell heterogeneity of HIV-infected cells during latent and activated stages, in a primary model of latently infected CD4+ T cells [58,60]. Using a virally-encoded fluorescent reporter gene to assess successful induction of HIV expression coupled with scRNA-Seq analysis, Golumbeanu *et al.* identified two cell subsets, with differential responsiveness to cell stimulation and HIV reactivation. One cell cluster displayed globally higher gene expression levels, and was more susceptible to cellular stimulation and HIV expression than the other cell cluster. Differential expression analysis also identified a 134 gene-specific signature to discriminate the two cell subsets.

Conclusions
Overall, scRNA-Seq is a powerful technique allowing characterizing cellular heterogeneity within a cell population. Association of transcriptome profilings with specific cellular phenotypes may help identifying specific proteins or biomarkers of interest involved in virus–host interactions, such as antibodies, TCR, virus receptor or cell-specific biomarkers.

Current analyses capture a unique snapshot of a single-cell gene expression profile at a specific time point, and do not allow manipulation of a single cell with follow-up analyses over time. Future technical improvements may possibly overcome this limitation, at least partially. Furthermore, development of novel protocols allowing simultaneous analysis of multiple molecular molecules, such as DNA, RNA and proteins, within the same individual cell would provide an additional step toward a more comprehensive 'single-cell integrome', which will likely be instrumental to further our understanding of viral replication and virus–host interactions.

Acknowledgements

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA *et al.*: **The sequence of**

the human genome. *Science (New York, NY)* 2001, **291**:1304-1351.

2. Manzoni C, Kia DA, Vandrovcova J, Hardy J, Wood NW, Lewis PA, Ferrari R: **Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences**. *Brief Bioinform* 2016:1-17 http://dx.doi.org/10.1093/bib/bbw114.

3. Haque A, Engel J, Teichmann SA, Lonnberg T: **A practical guide**
•• **to single-cell RNA-sequencing for biomedical research and clinical applications**. *Genome Med* 2017, **9**:75.
This review presents a practical guide to help researchers design a scRNA-Seq study. It provides a simple overview of the scRNA-Seq procedure as well as a comparison between the scRNA-Seq workflows available to date.

4. Su Z, Ning B, Fang H, Hong H, Perkins R, Tong W, Shi L: **Next-generation sequencing and its applications in molecular diagnostics**. *Exp Rev Mol Diagn* 2011, **11**:333-343.

5. Ohashi H, Hasegawa M, Wakimoto K, Miyamoto-Sato E: **Next-generation technologies for multiomics approaches including interactome sequencing**. *BioMed Res Int* 2015:104209.

6. Yoo BC, Kim KH, Woo SM, Myung JK: **Clinical multi-omics strategies for the effective cancer management**. *J Proteomics* 2017 http://dx.doi.org/10.1016/j.jprot.2017.08.010. [Epub ahead of print].

7. Wang H, Zhang Q, Fang X: **Transcriptomics and proteomics in stem cell research**. *Front Med* 2014, **8**:433-444.

8. Alessandro Tanca MD, Maria Filippa Addis, Sergio Uzzau: **High throughput genomic and proteomic technologies in the fight against infectious diseases**. *J Infect Dev Ctries* 2013, **7**:182-190.

9. Jean Beltran PM, Federspiel JD, Sheng X, Cristea IM: **Proteomics and integrative omic approaches for understanding host–pathogen interactions and infectious diseases**. *Mol Syst Biol* 2017, **13**:922.

10. Nagel ZD, Engelward BP, Brenner DJ, Begley TJ, Sobol RW, Bielas JH, Stambrook PJ, Wei Q, Hu JJ, Terry MB *et al.*: **Towards precision prevention: technologies for identifying healthy individuals with high risk of disease**. *Mutat Res* 2017, **800–802**:14-28.

11. Streeter OE, Beron PJ, Iyer PN: **Precision medicine: genomic profiles to individualize therapy**. *Otolaryngol Clin North Am* 2017, **50**:765-773.

12. Shalek AK, Benson M: **Single-cell analyses to tailor treatments**. *Sci Transl Med* 2017, **9**.

13. Altschuler SJ, Wu LF: **Cellular heterogeneity: do differences**
• **make a difference?** *Cell* 2010, **141**:559-563.
Global description of the concept of cell heterogeneity and the utility of single cell analysis.

14. Ciuffi A, Rato S, Telenti A: **Single-cell genomics for virology**.
•• *Viruses* 2016, **8**.
Overview of studies in virology using single-cell sequencing technologies.

15. Hrdlickova R, Toloue M, Tian B: **RNA-Seq methods for transcriptome analysis**. *Wiley Interdiscip Rev RNA* 2017, **8**.

16. Dagogo-Jack I, Shaw AT: **Tumour heterogeneity and resistance to cancer therapies**. *Nat Rev Clin Oncol* 2018, **15**:81-94 http://dx.doi.org/10.1038/nrclinonc.2017.166.

17. Wang Y, Navin NE: **Advances and applications of single-cell sequencing technologies**. *Mol Cell* 2015, **58**:598-609.

18. Chen Z, Chen L, Zhang W: **Tools for genomic and transcriptomic analysis of microbes at single-cell level**. *Front Microbiol* 2017, **8**:1831.

19. Vijayakumar S, Conway M, Lió P, Angione C: **Optimization of multi-omic genome-scale models: methodologies, hands-on tutorial, and perspectives**. In *Metabolic Network Reconstruction and Modeling: Methods and Protocols*. Edited by Fondi M. Springer; 2018:389-408.

20. Kim S, Kang D, Huo Z, Park Y, Tseng GC: **Meta-analytic principal component analysis in integrative omics application**. *Bioinformatics* 2017 http://dx.doi.org/10.1093/bioinformatics/btx765. [Epub ahead of print].

21. Czaplinski K:: **Techniques for single-molecule mRNA imaging in living cells**. In *Neuroepigenomics in Aging and Disease.* Edited by Delgado-Morales R. Springer International Publishing; 2017:425-441.

22. Stahlberg A, Kubista M: **Technical aspects and recommendations for single-cell qPCR**. *Mol Aspects Med* 2018, **59**:28-35 http://dx.doi.org/10.1016/j.mam.2017.07.004.

23. Hedlund E, Deng Q: **Single-cell RNA sequencing: technical advancements and biological applications**. *Mol Aspects Med* 2018, **59**:36-46 http://dx.doi.org/10.1016/j.mam.2017.07.003.

24. Baran-Gale J, Chandra T, Kirschner K: **Experimental design for single-cell RNA sequencing**. *Brief Funct Genom* 2017 http://dx.doi.org/10.1093/bfgp/elx035. [Epub ahead of print].

25. Ziegenhain C, Vieth B, Parekh S, Reinius B, Guillaumet-Adkins A, Smets M, Leonhardt H, Heyn H, Hellmann I, Enard W: **Comparative analysis of single-cell RNA sequencing methods**. *Mol Cell* 2017, **65** 631–643 e634.

26. Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, Satija R, Smibert P: **Simultaneous epitope and transcriptome measurement in single cells**. *Nat Methods* 2017, **14**:865-868.

27. Rato S, Golumbeanu M, Telenti A, Ciuffi A: **Exploring viral infection using single-cell sequencing**. *Virus Res* 2017, **239**:55-68.

28. Hu P, Zhang W, Xin H, Deng G: **Single cell isolation and analysis**. *Front Cell Dev Biol* 2016, **4**:116.

29. Gawad C, Koh W, Quake SR: **Single-cell genome sequencing: current state of the science**. *Nat Rev Genet* 2016, **17**:175-188.

30. Prakadan SM, Shalek AK, Weitz DA: **Scaling by shrinking:**
•• **empowering single-cell 'omics' with microfluidic devices**. *Nat Rev Genet* 2017, **18**:345-361.
Description of current state-of-the-art methodologies to isolate single cells using microfluidics for mammalian single-cell 'omics' analyses.

31. Fan X, Zhang X, Wu X, Guo H, Hu Y, Tang F, Huang Y: **Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos**. *Genome Biol* 2015, **16**:148.

32. Sheng K, Cao W, Niu Y, Deng Q, Zong C: **Effective detection of variation in single-cell transcriptomes using MATQ-seq**. *Nat Methods* 2017, **14**:267-270.

33. Islam S, Zeisel A, Joost S, La Manno G, Zajac P, Kasper M, Lonnerberg P, Linnarsson S: **Quantitative single-cell RNA-seq with unique molecular identifiers**. *Nat Methods* 2014, **11**:163-166.

34. Smith T, Heger A, Sudbery I: **UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy**. *Genome Res* 2017, **27**:491-499.

35. Grun D, van Oudenaarden A: **Design and analysis of single-cell sequencing experiments**. *Cell* 2015, **163**:799-810.

36. Kolodziejczyk AA, Kim JK, Svensson V, Marioni JC,
•• Teichmann SA: **The technology and biology of single-cell RNA sequencing**. *Mol Cell* 2015, **58**:610-620.
Nice and comprehensive review about the whole scRNA-Seq pipeline.

37. Picelli S: **Single-cell RNA-sequencing: the future of genome**
• **biology is now**. *RNA Biol* 2017, **14**:637-650.
Description of the diverse methods used to generate the scRNA-Seq libraries.

38. Rostom R, Svensson V, Teichmann SA, Kar G: **Computational**
• **approaches for interpreting scRNA-seq data**. *FEBS Lett* 2017, **591**:2213-2225.
The authors discuss a few examples of biological questions that can be investigated using scRNA-Seq approaches and the computational tools available to analyze and address them.

39. Stoeger T, Battich N, Pelkmans L: **Passive noise filtering by cellular compartmentalization**. *Cell* 2016, **164**:1151-1161.

40. Lau BT, Ji HP: **Single molecule counting and assessment of random molecular tagging errors with transposable giga-scale error-correcting barcodes**. *BMC Genom* 2017, **18**:745.

41. Audoux J, Philippe N, Chikhi R, Salson M, Gallopin M, Gabriel M, Le Coz J, Drouineau E, Commes T, Gautheret D: **DE-kupl: exhaustive capture of biological variation in RNA-seq data through k-mer decomposition**. *Genome Biol* 2017, **18**:243.

42. J.M.P.: **Single cell sequencing made simple**. *Nature* 2017, **547**:7661.

43. Bacher R, Kendziorski C: **Design and computational analysis of**
•• **single-cell RNA-sequencing experiments**. *Genome Biol* 2016, **17**:63.
Overview of computational methods available for the design and statistical analysis of scRNA-seq experiments, as well as their advantages and disadvantages.

44. Poirion OB, Zhu X, Ching T, Garmire L: **Single-cell transcriptomics bioinformatics and computational challenges**. *Front Genet* 2016, **7**:163.

45. Stegle O, Teichmann SA, Marioni JC: **Computational and**
•• **analytical challenges in single-cell transcriptomics**. *Nat Rev Genet* 2015, **16**:133-145.
Overview of computational challenges in scRNA-seq analyses and available bioinformatic tools to address them.

46. Yalcin D, Hakguder ZM, Otu HH: **Bioinformatics approaches to single-cell analysis in developmental biology**. *Mol Hum Reprod* 2016, **22**:182-192.

47. McWilliam Leitch EC, McLauchlan J: **Determining the cellular diversity of hepatitis C virus quasispecies by single-cell viral sequencing**. *J Virol* 2013, **87**:12648-12655.

48. Wu L, Zhang X, Zhao Z, Wang L, Li B, Li G, Dean M, Yu Q, Wang Y, Lin X *et al.*: **Full-length single-cell RNA-seq applied to a viral human cancer: applications to HPV expression and splicing analysis in HeLa S3 cells**. *GigaScience* 2015, **4**:51.

49. Tsioris K, Gupta NT, Ogunniyi AO, Zimnisky RM, Qian F, Yao Y, Wang X, Stern JN, Chari R, Briggs AW *et al.*: **Neutralizing antibodies against West Nile virus identified directly from human B cells by single-cell analysis and next generation sequencing**. *Integr Biol: Quant Biosci Nano Macro* 2015, **7**:1587-1597.

50. Ogunniyi AO, Story CM, Papa E, Guillen E, Love JC: **Screening individual hybridomas by microengraving to discover monoclonal antibodies**. *Nat Protoc* 2009, **4**:767-782.

51. Afik S, Yates KB, Bi K, Darko S, Godec J, Gerdemann U, Swadling L, Douek DC, Klenerman P, Barnes EJ *et al.*: **Targeted reconstruction of T cell receptor sequence from single cell RNA-seq links CDR3 length to T cell differentiation state**. *Nucleic acids Res* 2017, **45**:e148.

52. Nowakowski TJ, Pollen AA, Di Lullo E, Sandoval-Espinosa C, Bershteyn M, Kriegstein AR: **Expression analysis highlights AXL as a candidate Zika virus entry receptor in neural stem cells**. *Cell Stem Cell* 2016, **18**:591-596.

53. Onorati M, Li Z, Liu F, Sousa AMM, Nakagawa N, Li M, Dell'Anno MT, Gulden FO, Pochareddy S, Tebbenkamp ATN *et al.*: **Zika virus disrupts phospho-TBK1 localization and mitosis in human neuroepithelial stem cells and radial glia**. *Cell Rep* 2016, **16**:2576-2592.

54. Johnson TS, Abrams ZB, Mo X, Zhang Y, Huang K: **Lack of human cytomegalovirus expression in single cells from glioblastoma tumors and cell lines**. *J Neurovirol* 2017, **23**:671-678.

55. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, Cahill DP, Nahed BV, Curry WT, Martuza RL *et al.*: **Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma**. *Science (New York, NY)* 2014, **344**:1396-1401.

56. Martin-Gayo E, Cole M, Kolb KE, Ouyang Z, Kazer SW, Walker BD, Yosef N, Shalek AK, Yu XG: **Identification of a highly functional DC subset in controllers by single-cell RNA-Seq [CROI Abstract 19]**. *Special Issue: Abstracts From the 2016 Conference on Retroviruses and Opportunistic Infections. Boston (MA)*. USA: Topics in Antiviral Medicine; 2016. [vol. 24(e-1):7–8].

57. Rato S, Rausell A, Munoz M, Telenti A, Ciuffi A: **Single-cell analysis identifies cellular markers of the HIV permissive cell**. *PLoS Pathogens* 2017, **13**:e1006678.

58. Golumbeanu M, Rato S, Cristinelli S, Munoz M, Cavassini M, Beerenwinkel N, Ciuffi A: **Single-cell RNA-Seq reveals transcriptional heterogeneity in latent and reactivated HIV-infected cells (Abstract OP1.6)**. *Special Issue: Abstract From Eighth International Workshop on HIV Persistence during Therapy; Miami (FL), USA: 2017. [J Virus Eradic, vol. 3 (Suppl. 5): 3]*.

59. Martin-Gayo E, Yu XG: **Dendritic cell immune responses in HIV-1 controllers**. *Curr HIV/AIDS Rep* 2017, **14**:1-7.

60. Mohammadi P, di Iulio J, Munoz M, Martinez R, Bartha I, Cavassini M, Thorball C, Fellay J, Beerenwinkel N, Ciuffi A *et al.*: **Dynamics of HIV latency and reactivation in a primary CD4+ T cell model**. *PLoS Pathogens* 2014, **10**:e1004156.

61. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, Wang X,
•• Bodeau J, Tuch BB, Siddiqui A *et al.*: **mRNA-Seq whole-transcriptome analysis of a single cell**. *Nat Methods* 2009, **6**:377-382.
First study of the transcriptome of a single cell, implementing the first scRNA-Seq method.

62. Sasagawa Y, Nikaido I, Hayashi T, Danno H, Uno KD, Imai T, Ueda HR: **Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity**. *Genome Biol* 2013, **14**:R31.

63. Nakamura T, Yabuta Y, Okamoto I, Aramaki S, Yokobayashi S, Kurimoto K, Sekiguchi K, Nakagawa M, Yamamoto T, Saitou M: **SC3-seq: a method for highly parallel and quantitative measurement of single-cell gene expression**. *Nucleic Acids Res* 2015, **43**:e60.

64. Ramskold D, Luo S, Wang YC, Li R, Deng Q, Faridani OR, Daniels GA, Khrebtukova I, Loring JF, Laurent LC *et al.*: **Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells**. *Nat Biotechnol* 2012, **30**:777-782.

65. Picelli S, Bjorklund AK, Faridani OR, Sagasser S, Winberg G, Sandberg R: **Smart-seq2 for sensitive full-length transcriptome profiling in single cells**. *Nat Methods* 2013, **10**:1096-1098.

66. Islam S, Kjallquist U, Moliner A, Zajac P, Fan JB, Lonnerberg P, Linnarsson S: **Highly multiplexed and strand-specific single-cell RNA 5′ end sequencing**. *Nat Protoc* 2012, **7**:813-828.

67. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM *et al.*: **Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets**. *Cell* 2015, **161**:1202-1214.

68. Hashimshony T, Wagner F, Sher N, Yanai I: **CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification**. *Cell Rep* 2012, **2**:666-673.

69. Hashimshony T, Senderovich N, Avital G, Klochendler A, de Leeuw Y, Anavy L, Gennert D, Li S, Livak KJ, Rozenblatt-Rosen O *et al.*: **CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq**. *Genome Biol* 2016, **17**:77.

70. Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, Peshkin L, Weitz DA, Kirschner MW: **Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells**. *Cell* 2015, **161**:1187-1201.

71. Jaitin DA, Kenigsberg E, Keren-Shaul H, Elefant N, Paul F, Zaretsky I, Mildner A, Cohen N, Jung S, Tanay A *et al.*: **Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types**. *Science (New York, NY)* 2014, **343**:776-779.