Virus-inclusive single cell RNA sequencing reveals molecular signature predictive of progression to severe dengue infection

Fabio Zanini[1,*], Makeda L. Robinson[2,3*] , Derek Croote[1], Malaya Kumar Sahoo[4], Ana Maria Sanz[5], Eliana Ortiz-Lasso[6], Ludwig Luis Albornoz[6], Fernando Rosso Suarez[5,7], Jose G. Montoya[3], Benjamin A. Pinsky[3,4], Stephen R. Quake[1,8,9,+], Shirit Einav[2,3+]

*: equal contribution
+: to whom correspondence should be addressed: seinav@stanford.edu; quake@stanford.edu

[1]Department of Bioengineering, Stanford University, Stanford, California, USA

[2]Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, California, USA.

[3]Department of Medicine, Division of Infectious Diseases and Geographic Medicine, Stanford University School of Medicine, Stanford, California, USA.

[4]Department of Pathology, Stanford University School of Medicine, Stanford, California, USA

[5]Clinical Research Center, Fundación Valle del Lili, Cali, Colombia.

[6]Pathology and Laboratory Department, Fundación Valle del Lili, Cali, Colombia.

[7]Department of Internal Medicine, Division of Infectious Diseases, Fundación Valle del Lili, Cali, Colombia.

[8]Department of Applied Physics, Stanford University, Stanford, California, USA.

[9]Chan Zuckerberg Biohub, San Francisco, California, USA.

Keywords: dengue, single cell, transcriptomics, biomarkers, virus-host interactions

Abstract

Dengue virus (DENV) infection can result in severe complications. Yet, the understanding of the molecular correlates of severity is limited, partly due to difficulties in defining the peripheral blood mononuclear cells (PBMCs) that are associated with DENV in vivo. Additionally, there are currently no biomarkers predictive of progression to severe dengue (SD). Bulk transcriptomics data are difficult to interpret because blood consists of multiple cell types that may react differently to infection. Here we applied virus-inclusive single cell RNA-seq approach (viscRNA-Seq) to profile transcriptomes of thousands of single PBMCs derived early in the course of disease from six dengue patients and four healthy controls, and to characterize distinct DENV-associated leukocytes. Multiple genes, particularly interferon response genes, were upregulated in a cell-specific manner prior to progression to SD. Expression of MX2 in naive B cells and CD163 in CD14$^+$ CD16$^+$ monocytes was predictive of SD. The majority of DENV-associated cells in the blood of two patients who progressed to SD were naive IgM B cells expressing the CD69 and CXCR4 receptors and antiviral genes, followed by monocytes. Bystander uninfected B cells also demonstrated immune activation, and plasmablasts from two patients exhibited antibody lineages with convergently hypermutated heavy chain sequences. Lastly, assembly of the DENV genome revealed diversity at unexpected genomic sites. This study presents a multi-faceted molecular elucidation of natural dengue infection in humans and proposes biomarkers for prediction of SD, with implications for profiling any tissue and viral infection, and for the development of a dengue prognostic assay.

Significance

A fraction of the 400 million people infected with dengue annually progresses to severe dengue (SD). Yet, there are currently no biomarkers to effectively predict disease progression. We profiled the landscape of host transcripts and viral RNA in thousands of single blood cells from dengue patients prior to progressing to SD. We discovered cell-type specific immune activation and candidate predictive biomarkers. We also revealed preferential virus association with specific cell populations, particularly naive B cells and monocytes. We then explored immune activation of bystander cells, clonality and somatic evolution of adaptive immune repertoires, and viral genomics. This multi-faceted approach could advance understanding of pathogenesis of any viral infection, map an atlas of infected cells and promote the development of prognostics.

Introduction

Dengue virus (DENV) is a major threat to global health, estimated to infect 400 million people annually in over 100 countries [1]. The four serotypes of DENV are transmitted by a mosquito vector. A licensed vaccine has shown limited efficacy and increased hospitalizations in children

[2,3], and there are currently no approved antivirals available for dengue treatment. The majority of symptomatic patients present with dengue fever experiencing flu-like symptoms. Five to twenty percent of these patients progress to severe dengue (SD), manifested by bleeding, plasma leakage, shock, organ failure, and sometimes death [4,5]. Early administration of supportive care reduces mortality in patients with SD [6], however, there are no accurate means to predict which patients will progress to SD. The currently utilized warning signs to identify dengue patients at risk of progressing to severe disease are based on clinical parameters that appear late in the disease course and are neither sensitive nor specific. This promotes ineffective patient triage and resource allocation and continued morbidity and mortality [7–9].

DENV strains are classified in four clades called serotypes 1 to 4. The greatest risk factor for SD is previous infection with a heterologous DENV serotype, which can cause antibody-dependent enhancement (ADE) of the secondary infection [10–13]. Aberrant activation of cross-reactive T-cells may also play a role [14]. Biomarkers for early detection of SD based on molecular features of the patient's blood have been proposed. These efforts have focused on two experimental techniques: (i) flow cytometry of fixed blood cell populations [15], and (ii) gene expression in bulk RNA extracted from blood or peripheral mononuclear blood cells (PBMCs) [16–19]. Although useful, these studies suffer from several limitations. The majority of these studies identified genes whose altered expression is associated with but does not precede the onset of SD and therefore cannot be used as prognostic biomarkers. From a technical standpoint, flow cytometry has a high throughput but is constrained to a few protein markers that are selected a priori, making it excellent for separating known, discrete cell populations but less appropriate for screening the complex, dynamic landscape of cell types, subtypes, and states characteristic of immune responses. Transcriptomics performed on bulk cell populations can screen thousands of genes but its resolution is limited, because it cannot capture tissue heterogeneity. Averaging the signal over various cell populations is confounded by changes both in abundances of cell types and activation states. Coupling fluorescence activated cell sorting (FACS) with single cell transcriptomics can potentially combine the advantages of both approaches [20]. It has also been challenging to identify DENV-associated and DENV-infected immune cells in humans. Since the composition of the immune system is complex and dynamic, we hypothesized that an unbiased, whole-genome approach would be more suitable than techniques based on a small number of markers to pinpoint the exact subset of cells associated with DENV in vivo and identify biomarkers of severity in distinct cell populations.

To facilitate identification of host cells containing viral RNA and characterization of their transcriptional response, we recently developed a virus-inclusive single cell RNA-seq (viscRNA-seq) platform [21]. We previously used this platform to characterize transcriptome dynamics of infection with flaviviruses in cultured human hepatoma (Huh7) cells [21]. We simultaneously quantified host cell mRNA and virus RNA abundance at the single cell level, thereby identifying proviral and antiviral factors that correlated with intracellular viral abundance [21]. A similar approach was used to monitor host response to Zika virus infection in neuronal stem cells [22]. Here, we coupled FACS with viscRNA-Seq to identify virus-associated cells from human patients and studied the molecular signatures preceding the development of SD

infection. The use of antibodies against surface proteins during FACS enabled enrichment for specific cell populations. Moreover, since viscRNA-Seq requires no genetic manipulation of the cells of interest, this approach enabled high-resolution screening of the whole human transcriptome for changes in gene expression at the single cell level.

Results

High-dimensional profiling of single cells from dengue virus infected patients

We combined FACS with viscRNA-Seq to profile the host and viral transcriptomes in peripheral mononuclear blood cells (PBMCs) collected early in the course of natural dengue infection in humans. Blood samples were derived from four healthy control subjects and six DENV infected patients: two who experienced an uncomplicated disease course and four who subsequently progressed to SD (Fig. 1A-B). All subjects were prospectively enrolled to a cohort that we established in Colombia ("Colombia cohort") (Supplementary Tables 1 and 2). Disease severity was classified on-site using 2009 WHO criteria upon presentation and discharge [23]. Patients were enrolled within 2-5 days after symptoms onset based on clinical presentation compatible with dengue or dengue with warning signs and positive NS1 antigen and/or anti-DENV IgM antibody. Notably, patients who presented with SD at the earliest visit were excluded. Whole blood and serum samples were obtained upon presentation. qRT-PCR [24] and serological assays confirmed the diagnosis of DENV infection and excluded other arboviral infections (including Zika and chikungunya) [25]. IgG avidity testing distinguished primary from secondary dengue (Supplementary Table 1) [25]. PBMC samples were isolated, stored and shipped in liquid nitrogen.

To sort multiple types of immune cells in patient PBMC samples and enable viscRNA-Seq with high specificity and throughput, we assembled two panels of antibodies against host cell surface markers. The PBMC samples were split into several aliquots, immunostained (Fig. 1C and Supplementary Figures 1-3 and Supplementary Table 3), and sorted via FACS into T cells, natural killer (NK) cells, B cells, monocytes, and dendritic (DC) cells. The viscRNA-Seq protocol was then followed, and each cell was sequenced at a depth of ~1 million reads on NextSeq 500 and NovaSeq (illumina) instruments (Fig. 1D). To measure intracellular DENV RNA abundance, we conducted viscRNA-Seq using the previously reported pan-DENV capture oligo [21]. The information provided by this approach on each individual cell included the cell type, immune activation state, infection state (whether and how much DENV RNA the cell contains), and sequence of the virus strain (Fig. 1E).

FACS-assisted viscRNA-Seq captures multiple cell types and activation states

Most human tissues including blood present a skewed composition of cell types. Unbiased cell capture, as routinely done in microfluidics protocols (e.g. [26]), produces detailed data on the most abundant cell populations, but fails to represent rare cell populations. To overcome this limitation, we combined FACS with a plate-based protocol to capture immune cells from samples containing less than 1,000,000 cells (because cells are sorted directly into single wells)

with high sensitivity (as assessed by CD45 expression), and adequate representation of various cell populations (Fig. 2A) [27,28]. In total, we sequenced over 13,000 cells, of which several hundred showed robust signal for DENV RNA (Fig. 2A). Following quality filtering, tens to hundreds of cells were analyzed for most cell types of each sample, for a total of ~8,700 cells (Fig. 2B). Within each cell type, multiple distinct overlapping immune cell subtypes and cell states were well represented in the dataset (Fig. 2C and Supplementary Table 4). In particular, within B cells alone we profiled many naive, IgM/IGD double positive cells as well as isotype switched cells. Most B cells formed a continuum of differentiation, but we also identified two additional clusters: the first expressed markers of plasmablasts and plasma cells, whereas the second showed high expression of TYROBP, a transmembrane signalling protein that has been implicated in B cell proliferation [29].

Profiling single cell gene expression identifies candidate predictive biomarkers of severe dengue infection

Next, we profiled the host transcriptome responses in the various PBMC populations. Since blood samples were obtained early in the course of dengue infection, this analysis was aimed at revealing alterations in gene expression that preceded the progression to SD. For each cell subtype and gene, we compared the distribution of expression values across the three categories of subjects: healthy control (H); uncomplicated dengue (D), and severe dengue (SD). To identify differentially expressed genes, we used a two-sample Kolmogorov-Smirnov test together with a computation of fold change in the averages across cells. We identified several genes whose expression was strongly upregulated in subjects that subsequently progressed to SD. Many of these genes belonged to the antiviral interferon response, yet they were upregulated in a cell type specific manner (Fig. 3A). Some genes were expressed in multiple cell types but were upregulated more strongly in specific cells from SD subjects (Fig. 3B); other genes were expressed essentially only during SD except in a few cell types (Fig. 3C); a few genes were expressed only in one cell type and only in subjects who subsequently developed SD (e.g. CD163 in monocytes, Fig. 3D). These results indicate that distinct cell populations respond differently to the same viral infection, confounding the performance of bulk assays, such as microarrays. Since this heterogeneity is not a hindrance but rather a resource within the single cell approach, we then explored the predictive potential of gene expression in specific cell types. To do so, we averaged across cells within the same patient and cell population and tested binary classification of severity at increasing thresholds of expression, de facto simulating a pseudo-bulk assay that could be implemented in the clinic. We identified a number of genes in specific cell populations that showed great predictive power for distinguishing SD from other subjects, as assessed by receiver operating characteristic (ROC) curves (Fig. 3E). Two notable examples with optimal ROC performance (area under the curve = 1) are MX2 in naive B lymphocytes and CD163 in double positive CD14$^+$/CD16$^+$ monocytes.

Virus in severe dengue patients is primarily associated with naive B cells

To define the cell subtypes that are associated with DENV in our PBMC samples, we then focused on cells with viral RNA reads. We detected viral reads in two samples only (out of six

dengue confirmed samples analyzed), both of which were derived from subjects who had high viral loads in their serum and that subsequently progressed to SD (samples 1-026-1 and 1-036-1, Supplementary Table 2). In both samples, a small number of monocytes were associated with viral RNA. A weak upregulation of CD4 and other genes was observed in these virus-associated monocytes (Supplementary Fig. 4). The majority of virus-associated cells were B lymphocytes (Fig. 4A). No viral reads were detected in other types of leukocytes. These findings are in line with a previous report based on bulk qPCR assays [30]. The fraction of uniquely mapped reads corresponding to viral RNA in those cells was heterogeneous but generally 1% or less, corresponding to several hundred reads per cell but much lower than we measured in cultured Huh7 cells [21].

To determine whether a distinct subpopulation of B cells was specifically associated with DENV, we identified the most upregulated genes in the virus-associated population versus other B cells from the same patients. DENV-associated B cells were enriched but not exclusive to IgM/IGD isotypes as well as other markers of naive B lymphocytes, such as the transcription factor TCL1A (see Supplementary Fig. 5 for more markers) [31]. The surface receptors CD69, FCRL1, and CXCR4 that signal B cell activation and tissue-specific homing, and IRF1 that encodes an interferon related protein were also upregulated (Fig. 4B). We computed a 2-dimensional embedding of the B cells via t-Distributed Stochastic Neighbor Embedding (tSNE) from the two PBMC samples with detectable viral reads and measured no viral reads associated with cells belonging to the plasma cells or TYROBP$^+$ clusters (Fig. 4C and Supplementary Fig. 5) [32]. We then assembled the whole B cell receptor (BCR) locus *de novo* and found that virus-associated, IgM B cells tended to have less hypermutations than other IgM B cells from the same subjects (Fig. 4D). In contrast, V/J usage in heavy and light chains was not apparently different between virus-associated and bystander B cells in the same subjects (Supplementary Fig. 6). Moreover, the expression of a number of genes that we have shown correlated with intracellular viral abundance in DENV-infected Huh7 cells and are known to participate in intracellular viral dynamics was not altered in the virus-associated B cells, raising the possibility that these cells are associated but not infected with DENV (Supplementary Fig. 7). We performed single molecule fluorescence in situ hybridization (smFISH) to detect positive and negative strand DENV RNA from naive B cells and monocytes from a healthy blood donor. We observed no evidence of either DENV strand in these B cells, unlike monocytes or control Huh7 cells (Supplementary Fig. 8).

In addition to counting the DENV reads, we mapped them in an iterative manner and recovered ~300,000 viral reads from patient 1-026-1 and ~2,000 reads from patient 1-036-1. We obtained high coverage across the whole DENV genome and a third of the genome, respectively. The intrapatient population genomics showed a wide range of conservation levels, as determined by minor allele frequencies (Fig. 4E and Supplementary Fig. 9A). Site-specific Shannon entropy restricted to positions with 200 or more virus reads did not correlate with cross-sectional entropy in DENV serotype 3 (Fig. 4F and Supplementary Fig. 9B).

Hundreds of non DENV-associated B cells (henceforth called "bystanders") were recovered from samples containing DENV-associated cells. We computed differential gene expression between these bystanders and B cells from healthy controls and identified a strong antiviral response via interferon stimulated genes IFI6, IFI44L, and IFIT3 (Fig. 4G). Moreover, we considered whether the diversity of the immune repertoire (B and T cell receptors - BCR and TCR) could play a role in virus-cell association. Whereas assembled BCRs from patients with detected DENV-associated B cells scattered into small clones, the BCR repertoire of patients 1-013-1 and 1-020-1, who had no DENV-associated B cells, contained large clonal families comprised of multiple plasmablasts sharing similar antibody heavy chains, indicating a rapid and large clonal expansion in the B cell compartment (Fig. 4H). The fact that such plasmablast expansions were captured simply as part of these patients' circulating B cell populations was surprising given the vast diversity of possible BCR rearrangements [33] and could be indicative of a more extensive plasmablast response and concurrent rise in neutralizing antibody titers known to occur in response to acute dengue infection [34]. One clonal family had members belonging to both patients, while another featured two plasmablasts with nearly identical heavy chains, but distinct light chains, raising the possibility of parallel somatic evolution [35]. Within the T cell compartment, we found that clustering by TCRβ/δ CDR3s produced clonal families that were largely private to an individual, while clustering according to TCRα/γ CDR3s revealed known invariant T cell subsets, including invariant natural killer T cells (iNKT) and mucosal associated invariant T cells (MAIT), as well as public γ chain CDR3 sequences (Supplementary Fig. 10) [36].

Discussion

We recently developed viscRNA-Seq, a scalable approach to quantify host and non-polyadenylated viral RNAs from the same cell [21]. In the current study, we apply viscRNA-Seq to in vivo samples for the first time and show that it can be used to effectively profile the landscape of host transcripts and viral RNA in thousands of single immune cells during natural dengue infection of human subjects. The human samples studied here posed additional challenges, beyond those presented in cultured cells. First, the exact sequence of the viral strain infecting the patient was unknown. We therefore designed oligonucleotide for virus capture in a conserved region of the viral genome [21]. Second, the target cell types of DENV in vivo are incompletely characterized, mandating assembly of antibody panels for FACS that maximize the probability of capturing the virus-associated cell population(s) (Fig. 1C). Third, since cell viability and integrity of the RNA after freezing, shipping, and thawing the PBMC samples was much lower than in cultured cells, we increased the throughput and PCR preamplification to ensure a sufficient number of high-quality cells from each cell type. These modifications have successfully addressed these challenges, as indicated by our findings. Because the viscRNA-seq approach can be extended to any virus of interest and is compatible with surface markers for FACS, we expect it to be readily extendible to dissociated solid tissues, for instance to characterize viral reservoirs of various viral infections. A similar approach was recently applied to influenza infection in mouse lungs [37] and to in vitro Zika virus infection of neuronal stem cells [22].

There are currently no clinically usable biomarkers to predict the development of severe complications associated with DENV infection including bleeding, shock, vascular leakage, and organ failure [4]. Previous work on molecular biomarkers for dengue severity has focused on flow cytometry, which has high throughput but requires an a priori choice of a few tens of marker proteins [15], and bulk transcriptomics, which can quantify the expression of all genes in parallel but is confounded by the superimposition of cell types (e.g. B and T lymphocytes) and cell state (activation of SD specific genes) [16,17]. Since obtained prior to the progression to SD, the blood samples analyzed in this study, combined with the single cell resolution and ability to sample a wide range of cell types and activation states via viscRNA-seq, provided a unique opportunity to discover predictive biomarkers of severity. Our data suggests the expression of MX2 within naive B cells and of CD163 within CD14+ CD16+ monocytes is greaty upregulated prior to the development of SD (Fig. 3E). It has been previously reported that MX2 is one of only four interferon-induced genes induced in an IRF3 and IRF7 independent manner in DENV infected mice [38], and that CD163 in macrophages and CD14+ CD16+ monocytes contributes to the pathogenesis of SD [39,40]. Given the small number of subjects analyzed in this study, the predictive power of these candidate biomarkers warrants further validation in larger cohorts. Nevertheless, these findings underscore the utility of the viscRNA-Seq approach to identify prognostic biomarkers for dengue.

Cell lines such as Huh7 and primary cells such as monocyte-derived dendritic cells are commonly used to study DENV-host cell interactions, however, elucidating the association of DENV with human cells in vivo has proven challenging. To the best of our knowledge, this is the first study to perform high-dimensional profiling of DENV-associated cells in vivo. In agreement with the expectation that myeloid cells are more easily infected than other cell types [11], we found 21 monocytes containing viral RNA in two patients with high viral load prior to the progression to SD. We observed a weak upregulation of CD4 and other genes in these virus-associated monocytes (Supplementary Fig. 4). Many more B cells from the same patients contained viral RNA, in line with a previous report on bulk samples [30]. Distribution-level statistics and dimensionality reduction indicated that IgM positive, naive B cells expressing the surface markers CD69 and CXCR4 are the subset of PBMCs that most commonly associate with DENV in vivo, although other B cells of other isotypes were also represented. In contrast, plasmablasts, plasma cells, and TYROBP positive B cells had no viral reads (Fig. 4A-D). B cells are known to be involved in SD pathogenesis by producing antibodies for antibody dependent enhancement (ADE) [12–14]. Discovering that naive B cells possessing diverse B cell receptor (BCR) sequences, rather than specific isotype switched, sequence-restricted memory B cells, are associated with DENV was therefore surprising. More work is needed to conclusively define whether these B cells are infected with DENV or just associated with it. Nevertheless, our data thus far based on lack of alteration in expression of genes known to be altered in DENV infection and the absence of DENV positive B cells in PBMCs via FISH support the latter. It is possible that these virus-associated B cells shuttle DENV within the human body. Bystander B cells from patients with detected DENV-associated cells had elevated levels of immune genes, particularly of the interferon response (Fig. 4G). B cells from one severe dengue and one

dengue patient samples (1-013-1 and 1-020-1) showed an interesting clonal structure in terms of antibody sequences. Specifically, multiple antibody sequences of heavy and light chains from several cells from patient 1-013 clustered into a few, presumably very large lineages of mostly heavily hypermutated IgG1 plasmablasts. Since no DENV RNA reads were detected in these samples (in contrast to samples 1-026-1 and 1-036-1), we speculate that this oligoclonal plasmablast population may be reducing binding of DENV by the host B cells. Further work is needed to understand the role of virus-associated and bystander B cells, determine whether DENV prevents the hypermutation upon binding naive B cells, and decipher whether the detected large antibody lineages are protective against DENV itself.

From the viral reads of two patients, 1-026-1 and 1-036-1, we assembled the entire or a third of the DENV genome, respectively, and observed some high-variability genomic sites (Fig. 4E). Previous work on other RNA viruses, particularly HIV-1, has shown that due to error-prone viral polymerases and fast generation times, intrapatient genomic viral diversity can represent a subsampled snapshot of the global diversity of the same virus in multiple infected individuals, implying a universal landscape of fitness costs [41,42]. DENV behaves quite differently, as globally variable sites do not correspond to variable sites within our patients (Fig. 4F). An optimized approach with higher sensitivity and sample selection (PBMCs or solid tissues) that maximizes the number of viral reads will facilitate a deeper understanding of the genomic diversity of viruses inhabiting the human body at the single cell level.
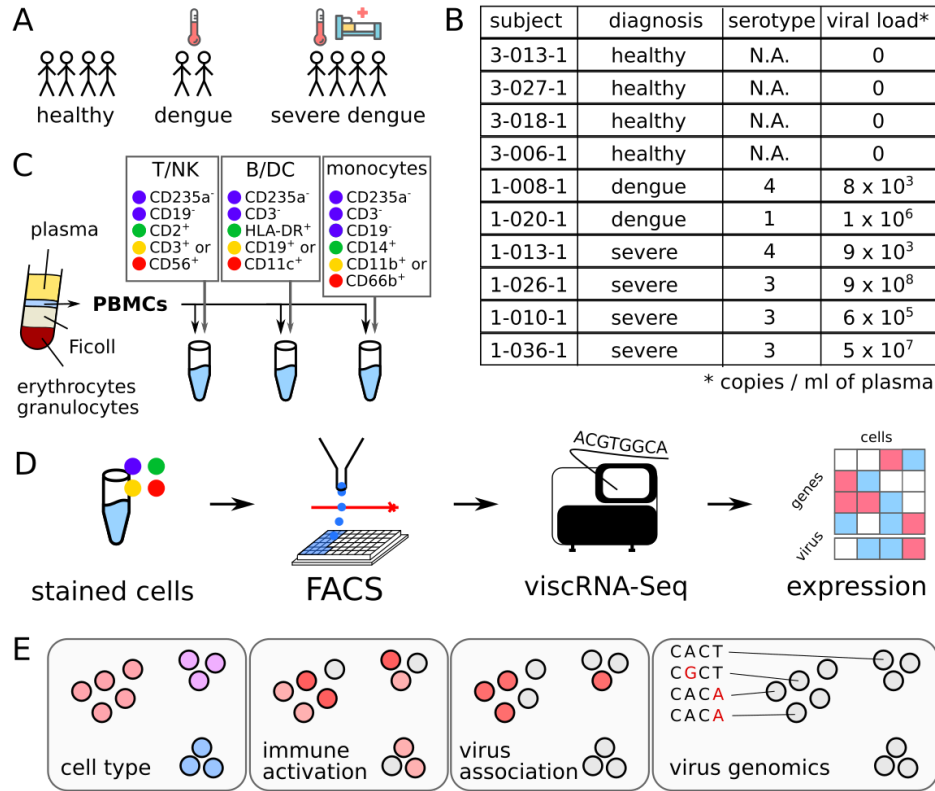
In this study, we leveraged the viscRNA-Seq approach to explore many different facets of virus infection in uncomplicated and severe dengue in humans at the single cell level. This multi-faceted profiling included investigation of transcriptional upregulation in specific subpopulations as a predictor of disease severity. Further validation in larger cohorts is warranted to determine the effectiveness of the identified candidate biomarkers as potential prognostic tools. Cell purification (e.g. by magnetic beads) followed by a rapid bulk expression assay (e.g. qPCR) is one option to translate such findings into a near-care, sample-to-answer system assay to be used for predicting progression of SD upon patient presentation. We also explored preferential association of virus with certain host cells, immune activation of bystander cells, clonality and somatic evolution of the adaptive immune repertoire, and intrapatient viral genomics. This technological convergence, combined with a high level of experimental and computational automation, underscores the utility of viscRNA-Seq as a powerful tool to rapidly gain a broad knowledge of emerging infectious diseases from just a few tissue samples.

Acknowledgements

Figures

Figure 1: Fluorescence-assisted viscRNA-Seq workflow on peripheral mononuclear blood cells from DENV infected and healthy control human subjects. (A-B) Blood samples were collected from human subjects enrolled to the Colombia cohort (healthy, dengue, and severe dengue). (C) Peripheral blood mononuclear cells (PBMCs) were isolated via Ficoll centrifugation and stained with three antibody panels to distinguish various cell types: T, B, natural killer (NK), dendritic cells (DC), and monocytes. (D) Single cells from each aliquot were sorted and processed by viscRNA-Seq to simultaneously quantify single cell virus abundance and host transcriptome changes. (E) The information provided for each single cell includes: cell type, immune activation, infection state, and virus population genomics. N.A.= non-applicable.
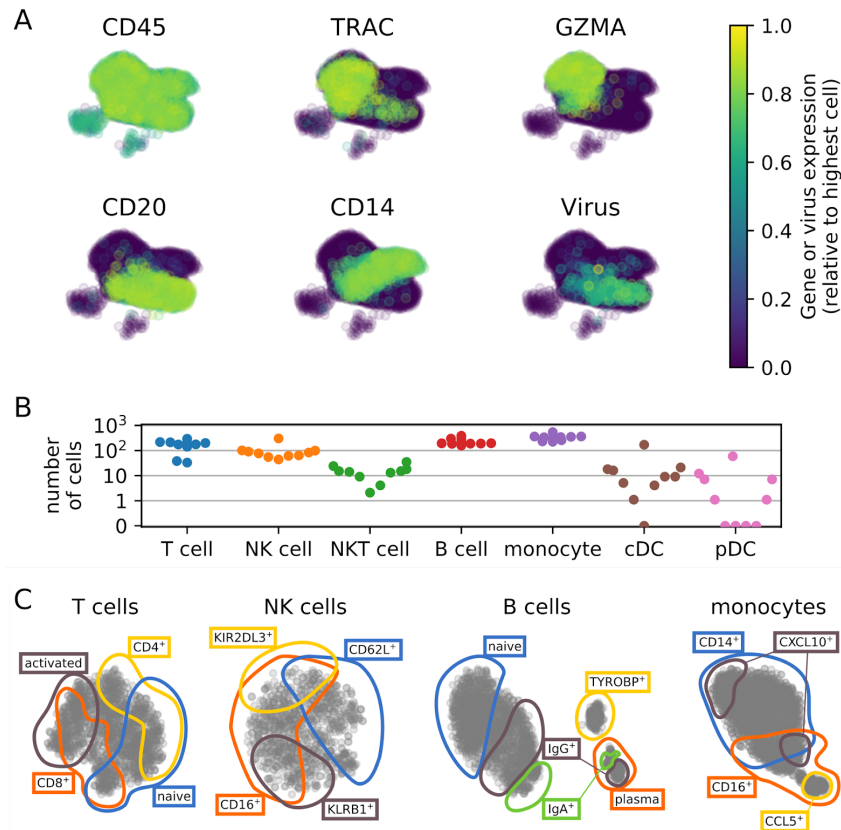
Figure 2: Overview of the types of peripheral mononuclear blood cells (PBMCs) surveyed. (A) Two-dimensional representation of the cells color coded by the expression level of cell type specific marker genes or the abundance of virus reads within the cell (> 30 virus reads per million reads in samples from two severe dengue patients, p1-026-1 and p1-036-1). (B) Number of cells analyzed for each cell type from each subject. (C) t-Distributed Stochastic Neighbor Embedding (tSNE) visualizations within T, NK, B cells, and monocytes, highlighting broad cell subtypes.
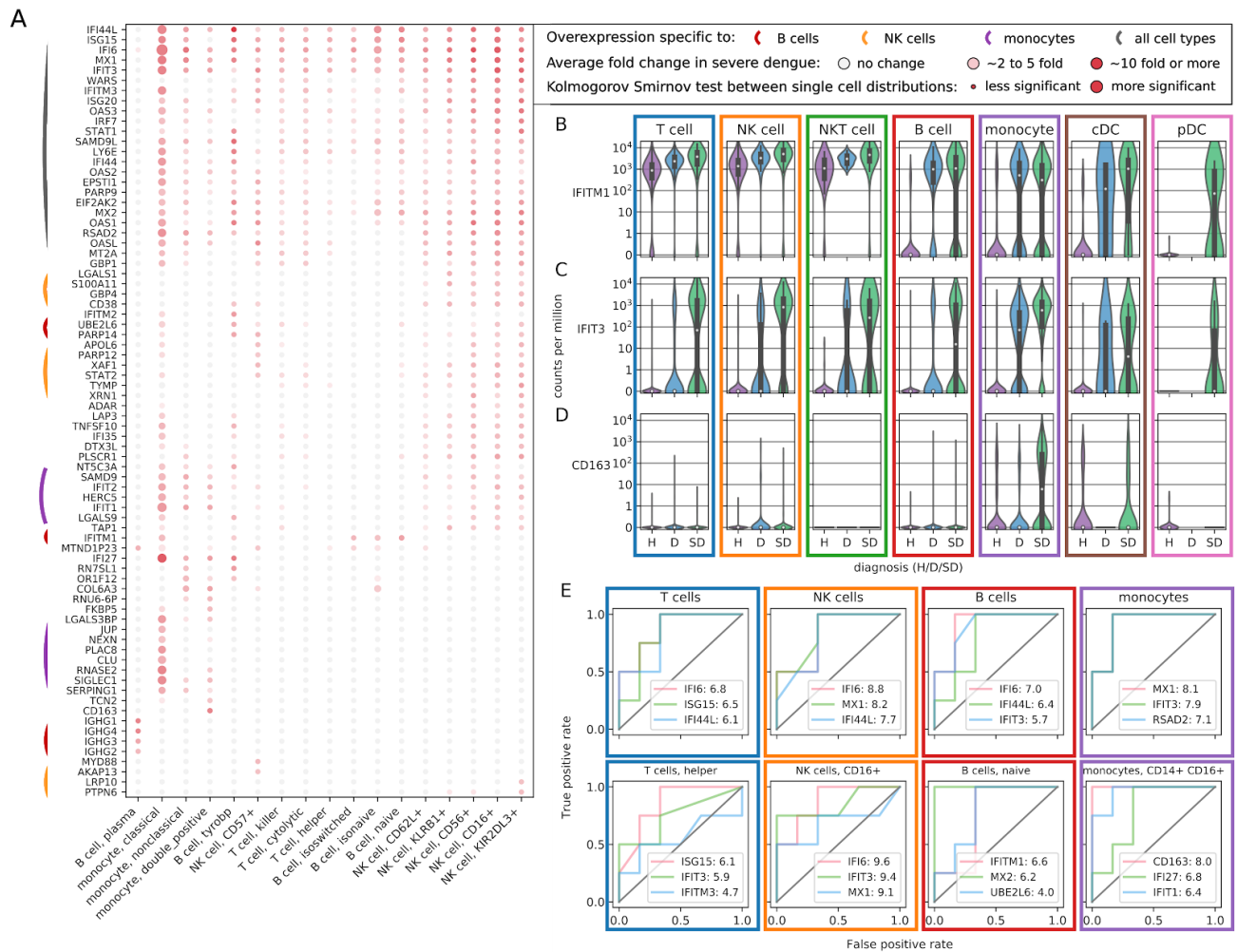
Figure 3: Differential expression across disease severity and cell types shows hallmarks predictive of severe dengue. (A) Genes that are overexpressed in subjects prior to progressing to severe dengue across cell types and subtypes. Color (white to red) indicates the average logfold change; size of the dot indicates lower P-value in a distribution statistical comparison (2 sample Kolmogorov-Smirnov). (B) Many inflammatory genes such as IFITM1 are expressed ubiquitously during both mild and severe dengue infection. (C) Other genes such as IFIT3 are specifically expressed prior to the development of severe dengue in various types of lymphocytes. (D) A number of genes show double specificity for both severe dengue and a single cell type, for instance CD163 in monocytes. (E) Averaging across cells within specific cell types and subtypes indicates promising candidate predictors of severe dengue as assessed by ROC curves at increasing discriminatory thresholds for gene expression versus disease severity. The numbers after the gene name indicate log2 fold changes of average expression in patients progressing to severe dengue versus other dengue patients, indicating an overexpression of these genes by a hundred fold or more in our cohort. H: healthy subject, D: dengue, SD: severe dengue.
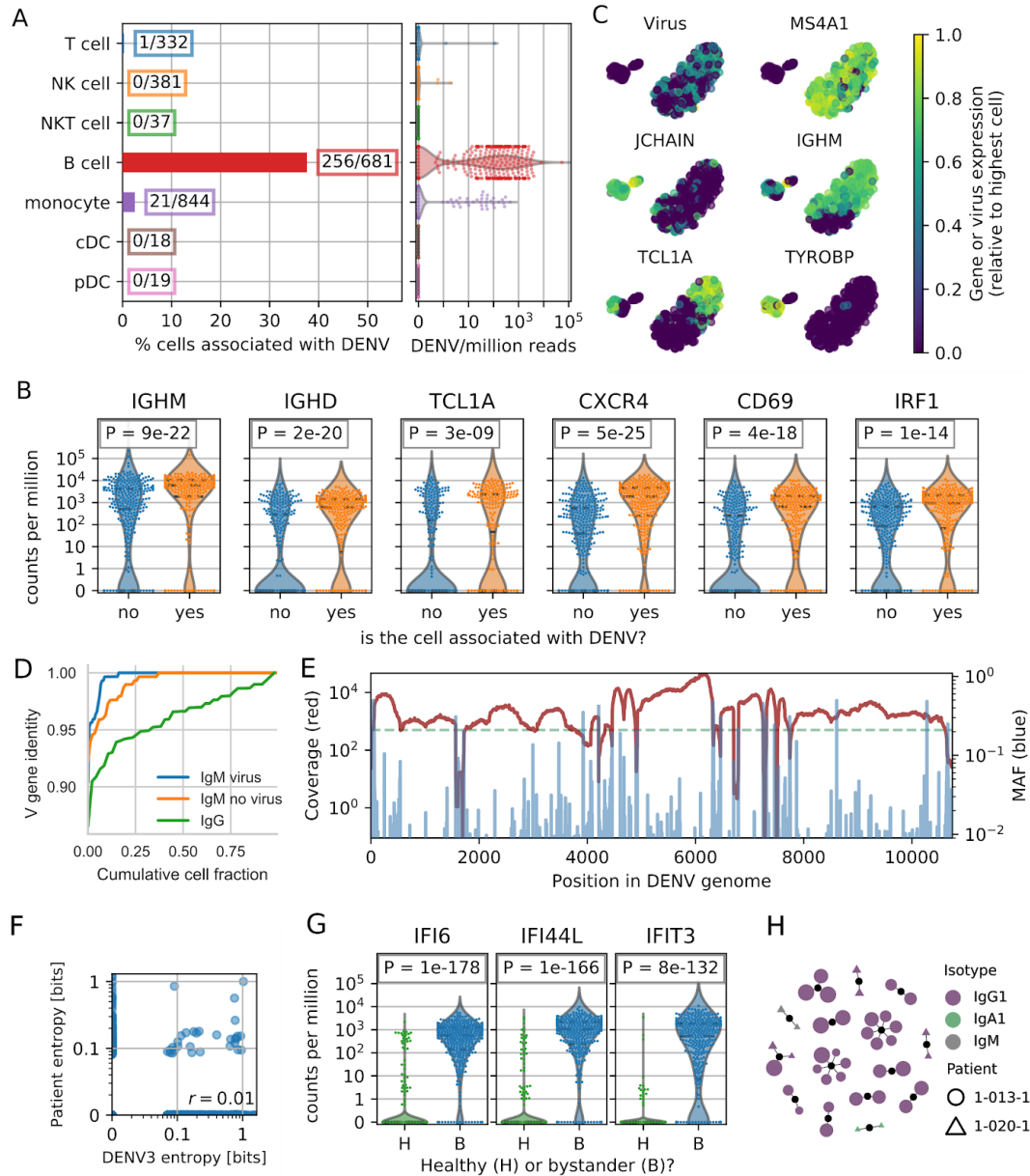
Figure 4: Dengue virus in two severe dengue patients is mostly associated with naive B cells. (A) Fraction of DENV-associated cells across cell types from the two subjects and relative amount of virus RNA from each cell. (B) Virus-associated B cells from the same subjects show a higher expression of specific surface receptors (CXCR4, CD69) and immune activation genes (IRF1, FCRL1). (C) tSNE visualization of the B cells from the two subjects. The expression level of DENV RNA and MS4A1 (CD20), JCHAIN, IGHM, TCL1A, and TYROBP are highlighted. (D) Fractional identity of heavy chain V loci to their germline counterparts in virus-associated IgM, bystander IgM, and IgG B cells from the subjects 1-026-1 and 1-036-1. (E) Coverage (red) and minor allele frequency (MAF, blue) along the DENV genome in the viral reads from all cells from patient sample 1-026-1 show the genetic diversity of the virus population. (F) Site specific Shannon entropy of a cross-sectional DENV serotype 3 alignment does not correlate with entropy from the viral reads of patient sample 1-026-1. Only sites with a coverage of 500 or

more reads are considered (dashed green line in panel E). (G) B cells that are not associated with DENV (bystanders) but derived from subjects with virus-associated cells (B, blue) show a clear interferon response compared with B cells derived from healthy controls (H, green). (H) Graph of heavy chain CDR3 antibody clonality showing clonal expansion of IgG1 plasmablasts in patients 1-013-1 and 1-020-1. Each dot is a unique antibody sequence, larger size corresponds to more somatic hypermutation.

References

1. Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, et al. The global distribution and burden of dengue. Nature. 2013;496: 504–507.

2. Guzman MG, Kouri G. Dengue and dengue hemorrhagic fever in the Americas: lessons and challenges. J Clin Virol. 2003;27: 1–13.

3. Khursheed M, Khan UR, Ejaz K, Fayyaz J, Qamar I, Razzak JA. A comparison of WHO guidelines issued in 1997 and 2009 for dengue fever - single centre experience. J Pak Med Assoc. 2013;63: 670–674.

4. WHO | Dengue guidelines for diagnosis, treatment, prevention and control: new edition. World Health Organization; 2011; Available: http://www.who.int/rpc/guidelines/9789241547871/en/

5. Dengue and severe dengue. In: World Health Organization [Internet]. 2018 [cited 25 Jun 2018]. Available: http://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue

6. Organization WH, Unicef, Others. Handbook for clinical management of dengue. World Health Organization; 2012; Available: http://apps.who.int/iris/bitstream/handle/10665/76887/?sequence=1

7. Hadinegoro SRS. The revised WHO dengue case classification: does the system need to be modified? Paediatr Int Child Health. 2012;32 Suppl 1: 33–38.

8. Nujum ZT, Thomas A, Vijayakumar K, Nair RR, Pillai MR, Indu PS, et al. Comparative performance of the probable case definitions of dengue by WHO (2009) and the WHO-SEAR expert group (2011). Pathog Glob Health. 2014;108: 103–110.

9. Thein TL, Leo Y-S, Lee VJ, Sun Y, Lye DC. Validation of probability equation and decision tree in predicting subsequent dengue hemorrhagic fever in adult dengue inpatients in Singapore. Am J Trop Med Hyg. 2011;85: 942–945.

10. Halstead SB, Chow JS, Marchette NJ. Immunological enhancement of dengue virus replication. Nat New Biol. 1973;243: 24–26.

11. Halstead SB, O'Rourke EJ. Dengue viruses and mononuclear phagocytes. I. Infection enhancement by non-neutralizing antibody. J Exp Med. 1977;146: 201–217.

12. Katzelnick LC, Gresh L, Halloran ME, Mercado JC, Kuan G, Gordon A, et al. Antibody-dependent enhancement of severe dengue disease in humans. Science. 2017;358: 929–932.

13. Wang TT, Sewatanon J, Memoli MJ, Wrammert J, Bournazos S, Bhaumik SK, et al. IgG antibodies to dengue enhanced for FcγRIIIA binding determine disease severity. Science. 2017;355: 395–398.

14. Whitehorn J, Simmons CP. The pathogenesis of dengue. Vaccine. 2011;29: 7221–7228.

15. Durbin AP, Vargas MJ, Wanionek K, Hammond SN, Gordon A, Rocha C, et al. Phenotyping of peripheral blood mononuclear cells during acute dengue illness demonstrates infection and increased activation of monocytes in severe cases compared to classic dengue fever. Virology. 2008;376: 429–435.

16. Ubol S, Masrinoul P, Chaijaruwanich J, Kalayanarooj S, Charoensirisuthikul T, Kasisith J. Differences in Global Gene Expression in Peripheral Blood Mononuclear Cells Indicate a Significant Role of the Innate Responses in Progression of Dengue Fever but Not Dengue Hemorrhagic Fever. J Infect Dis. 2008;197: 1459–1467.

17. Fink J, Gu F, Ling L, Tolfvenstam T, Olfat F, Chin KC, et al. Host Gene Expression Profiling of Dengue Virus Infection in Cell Lines and Patients. PLoS Negl Trop Dis. 2007;1: e86.

18. Sessions OM, Tan Y, Goh KC, Liu Y, Tan P, Rozen S, et al. Host Cell Transcriptome Profile during Wild-Type and Attenuated Dengue Virus Infection. PLoS Negl Trop Dis. 2013;7: e2107.

19. Nikolayeva I, Bost P, Casademont I, Duong V, Koeth F, Prot M, et al. A Blood RNA Signature Detecting Severe Disease in Young Dengue Patients at Hospital Arrival. J Infect Dis. 2018;217: 1690–1698.

20. Darmanis S, Sloan SA, Croote D, Mignardi M, Chernikova S, Samghababi P, et al. Single-Cell RNA-Seq Analysis of Infiltrating Neoplastic Cells at the Migrating Front of Human Glioblastoma. Cell Rep. 2017;21: 1399–1410.

21. Zanini F, Pu S-Y, Bekerman E, Einav S, Quake SR. Single-cell transcriptional dynamics of flavivirus infection. Elife. 2018;7. doi:10.7554/eLife.32942

22. Gorman MJ, Caine EA, Zaitsev K, Begley MC, Weger-Lucarelli J, Uccellini MB, et al. An Immunocompetent Mouse Model of Zika Virus Infection. Cell Host Microbe. 2018;23: 672–685.e6.

23. World Health Organization. Dengue: Guidelines for Diagnosis, Treatment, Prevention and Control. World Health Organization; 2009.

24. Waggoner JJ, Abeynayake J, Sahoo MK, Gresh L, Tellez Y, Gonzalez K, et al. Single-Reaction, Multiplex, Real-Time RT-PCR for the Detection, Quantitation, and Serotyping of Dengue Viruses. PLoS Negl Trop Dis. 2013;7: e2116.

25. Zhang B, Pinsky BA, Ananta JS, Zhao S, Arulkumar S, Wan H, et al. Diagnosis of Zika virus infection on a nanotechnology platform. Nat Med. 2017;23: 548–550.

26. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell. 2015;161: 1202–1214.

27. Zanini F. lshknn [Internet]. Stanford University; 2018. Available: https://github.com/iosonofabio/lshknn

28. Carnevali P. ExpressionMatrix2 [Internet]. Chan Zuckerberg Initiative; 2018. Available: https://github.com/chanzuckerberg/ExpressionMatrix2

29. Nakano-Yokomizo T, Tahara-Hanaoka S, Nakahashi-Oda C, Nabekura T, Tchao NK, Kadosaki M, et al. The immunoreceptor adapter protein DAP12 suppresses B lymphocyte-driven adaptive immune responses. J Exp Med. 2011;208: 1661–1671.

30. Srikiatkhachorn A, Wichit S, Gibbons RV, Green S, Libraty DH, Endy TP, et al. Dengue viral RNA levels in peripheral blood mononuclear cells are associated with disease severity and preexisting dengue immune status. PLoS One. 2012;7: e51335.

31. Said JW, Hoyer KK, French SW, Rosenfelt L, Garcia-Lloret M, Koh PJ, et al. TCL1 oncogene expression in B cell subsets from lymphoid hyperplasia and distinct classes of B cell lymphoma. Lab Invest. 2001;81: 555–564.

32. Maaten L van der, Hinton G. Visualizing Data using t-SNE. J Mach Learn Res. 2008;9: 2579–2605.

33. Georgiou G, Ippolito GC, Beausang J, Busse CE, Wardemann H, Quake SR. The promise and challenge of high-throughput sequencing of the antibody repertoire. Nat Biotechnol. 2014;32: 158–168.

34. Appanna R, Kg S, Xu MH, Toh Y-X, Velumani S, Carbajo D, et al. Plasmablasts During Acute Dengue Infection Represent a Small Subset of a Broader Virus-specific Memory B Cell Pool. EBioMedicine. 2016;12: 178–188.

35. Parameswaran P, Liu Y, Roskin KM, Jackson KKL, Dixit VP, Lee J-Y, et al. Convergent Antibody Signatures in Human Dengue. Cell Host Microbe. 2013;13: 691–700.

36. Ravens S, Schultze-Florey C, Raha S, Sandrock I, Drenker M, Oberdörfer L, et al. Human γδ T cells are quickly reconstituted after stem-cell transplantation and show adaptive clonal expansion in response to viral infection. Nat Immunol. 2017;18: 393–401.

37. Steuerman Y, Cohen M, Peshes-Yaloz N, Valadarsky L, Cohn O, David E, et al. Dissection of Influenza Infection In Vivo by Single-Cell RNA Sequencing. Cell Syst. 2018;6: 679–691.e4.

38. Chen H-W, King K, Tu J, Sanchez M, Luster AD, Shresta S. The roles of IRF-3 and IRF-7 in innate antiviral immunity against dengue virus. J Immunol. 2013;191: 4194–4201.

39. Kwissa M, Nakaya HI, Onlamoon N, Wrammert J, Villinger F, Perng GC, et al. Dengue virus infection induces expansion of a CD14(+)CD16(+) monocyte population that stimulates plasmablast differentiation. Cell Host Microbe. 2014;16: 115–127.

40. Ab-Rahman HA, Rahim H, AbuBakar S, Wong P-F. Macrophage Activation Syndrome-Associated Markers in Severe Dengue. Int J Med Sci. 2016;13: 179–186.

41. Zanini F, Brodin J, Thebo L, Lanz C, Bratt G, Albert J, et al. Population genomics of intrapatient HIV-1 evolution. Elife. 2016;4: e11282.

42. Zanini F, Puller V, Brodin J, Albert J, Neher RA. In vivo mutation rates and the landscape of

fitness costs of HIV-1. Virus Evolution. 2017;3. doi:10.1093/ve/vex003

43. Alexander N, Balmaseda A, Coelho ICB, Dimaano E, Hien TT, Hung NT, et al. Multicentre prospective study on dengue classification in four South-east Asian and three Latin American countries. Trop Med Int Health. 2011;16: 936–948.

44. Waggoner JJ, Gresh L, Mohamed-Hadley A, Ballesteros G, Davila MJV, Tellez Y, et al. Single-Reaction Multiplex Reverse Transcription PCR for Detection of Zika, Chikungunya, and Dengue Viruses. Emerg Infect Dis. 2016;22: 1295–1297.

45. Quake SR, Wyss-Coray T, Darmanis S. Single-cell transcriptomic characterization of 20 organs and tissues from individual mice creates a Tabula Muris. bioRxiv. biorxiv.org; 2018; Available: https://www.biorxiv.org/content/early/2018/03/29/237446.abstract

46. van der Walt S, Colbert SC, Varoquaux G. The NumPy Array: A Structure for Efficient Numerical Computation. Computing in Science Engineering. 2011;13: 22–30.

47. McKinney W. pandas: a Foundational Python Library for Data Analysis and Statistics. 2011;

48. Hunter JD. Matplotlib: A 2D graphics environment. Comput Sci Eng. 2007;9: 90–95.

49. Waskom M, Botvinnik O, Hobson P, Cole JB, Halchenko Y, Hoyer S, et al. seaborn: v0.5.0 (November 2014) [Internet]. 2014. doi:10.5281/zenodo.12710

50. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. J Mach Learn Res. 2011;12: 2825–2830.

51. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29: 15–21.

52. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31: 166–169.

53. Lunter G, Goodson M. Stampy: A statistical algorithm for sensitive and fast mapping of Illumina sequence reads. Genome Res. 2011;21: 936–939.

54. Kurtzer GM, Sochat V, Bauer MW. Singularity: Scientific containers for mobility of compute. PLoS One. 2017;12: e0177459.

55. Yang X, Charlebois P, Gnerre S, Coole MG, Lennon NJ, Levin JZ, et al. De novo assembly of highly diverse viral populations. BMC Genomics. 2012;13: 475.

56. Canzar S, Neu KE, Tang Q, Wilson PC, Khan AA. BASIC: BCR assembly from single cells. Bioinformatics. 2017;33: 425–427.

57. Ye J, Ma N, Madden TL, Ostell JM. IgBLAST: an immunoglobulin variable domain sequence analysis tool. Nucleic Acids Res. 2013;41: W34–40.

58. Gupta NT, Heiden JAV, Uduman M, Gadala-Maria D, Yaari G, Kleinstein SH. Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data.

Bioinformatics. 2015;31: 3356–3358.

59. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10: 421.

60. Lefranc M-P, Giudicelli V, Ginestoux C, Jabado-Michaloud J, Folch G, Bellahcene F, et al. IMGT, the international ImMunoGeneTics information system. Nucleic Acids Res. 2009;37: D1006–12.

61. Peixoto TP. The graph-tool python library [Internet]. 2017. doi:10.6084/m9.figshare.1164194.v14