No cell is an island: circulating T cell:monocyte complexes are markers of immune perturbations

Authors: Julie G. Burel[1], Mikhail Pomaznoy[1], Cecilia S. Lindestam Arlehamn[1], Daniela Weiskopf[1], Ricardo da Silva Antunes[1], Veronique Schulten[1], Mariana Babor[1], Grégory Seumois[1], Jason A. Greenbaum[1], Sunil Premawansa[2], Gayani Premawansa[3], Ananda Wijewickrama[4], Rashmi Tippalagama[5], Aruna D. deSilva[1,5], Robert H. Gilman[6,7], Mayuko Saito[8], Randy Taplitz[9], Pandurangan Vijayanand[1,10], Alessandro Sette[1,10], Bjoern Peters[1,10,*]

Affiliations:

[1]Division of Vaccine Discovery, La Jolla Institute for Allergy and Immunology, La Jolla, CA, USA.

[2]Department of Zoology and Environment Science, Science Faculty, University of Colombo, Sri Lanka.

[3]North Colombo Teaching Hospital, Ragama, Sri Lanka.

[4]National Institute of Infectious Diseases, Gothatuwa, Angoda, Sri Lanka.

[5]Genetech Research Institute, Colombo, Sri Lanka #present address: Dept of Paraclinical Sciences, Faculty of Medicine, General Sir John Kotelawala Defence University, Ratmalana, Sri Lanka.

[6]Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD, USA.

[7]Universidad Peruana Caytano Hereida, Lima, Peru.

[8]Department of Virology, Tohoku University Graduate School of Medicine, Sendai, Japan.

[9]Division of infectious diseases, University of California San Diego, La Jolla, CA, USA.

[10]Department of Medicine, University of California San Diego, La Jolla, CA, USA.

*Correspondence to: bpeters@lji.org

Abstract: We set out to detect a transcriptional signature in CD4 T cells of individuals at risk of progression to active tuberculosis. We indeed found such a signature which surprisingly could be tracked to a cell population expressing both T cell (CD3) and monocyte (CD14) markers. Imaging experiments revealed that $CD3^{+}CD14^{+}$ cells included T cell:monocyte complexes with high TCR signaling activity. Broadening our investigation beyond tuberculosis, we also found $CD3^{+}CD14^{+}$-complexes as a function of time post tetanus, diphtheria and pertussis (Tdap) vaccination and as a function of disease severity in dengue infection. These data suggest that circulating Tcell:monocyte complexes form *in vivo* as a result of T cells recognizing antigen presented on the monocytes, and that the conventional wisdom in flow cytometry to avoid studying cell:cell complexes should be re-visited.

One sentence summary: T cell:monocyte complexes observed in flow cytometry are not a technical artefact but reflect true biological interaction between immune cells.

42 **Introduction**

43 Detecting immune perturbations can help diagnose diseases. Our specific interest has been in
44 Tuberculosis (TB), a highly prevalent disease caused by the bacterial pathogen *Mycobacterium*
45 *tuberculosis* (Mtb) with nearly 2 billion people affected worldwide [1]. TB infections form a
46 spectrum: on one end, there are individuals with controlled infection (latent tuberculosis infection,
47 or LTBI) and on the other end individuals with active disease [2]. Only a small fraction (5-15%)
48 of individuals with LTBI will ever develop active disease, which is associated with significant
49 morbidity and mortality [1]. Treatment for LTBI individuals to eradicate the bacteria is an effective
50 way to prevent active disease. However, due to the risk of side-effects associated with these
51 treatments and low proportion of individuals actually progressing to active disease, treatment of
52 all LTBI individuals to eliminate the disease in high prevalence countries is not justified.

53 To overcome this hurdle, the WHO has called for the development of diagnostic tests that can
54 reliably predict LTBI progressors with a sensitivity and specificity of 90% [3]. This would enable
55 targeted treatment of at-risk LTBI individuals with the goal of eradicating the disease. Recent
56 efforts towards the development of such tests has included transcriptomic profiling which
57 discovered concise gene signatures predictive of LTBI progression [4, 5]. These results
58 demonstrate the applicability of transcriptional profiling to develop diagnostic tests for risk of TB
59 progression, but the performance of these tests is still far from reaching WHO's goal, and there is
60 still no mechanistic understanding of immune responses required for protection.

61 We recently identified a gene signature in memory $CD4^+$ T cells that distinguishes individuals with
62 LTBI from uninfected individuals [6]. Such transcriptional profiling of selected cell populations
63 is not a realistic option as a diagnostic test for TB in low resource environments, but it can provide
64 a clearer picture of the role that $CD4^+$ T cells play in TB immunity and in the mechanisms of

65 disease progression, which is harder to do in whole blood studies [7]. We targeted $CD4^+$ T cells
66 as they are known to be key players in TB protective immunity [8, 9]. In particular, antigen-specific
67 $CD4^+$ T cell responses have been shown to be associated with protection in the context of BCG
68 vaccination [10], and defects in the development of $CD4^+CCR6^+CXCR3^+$ T cells, which contain
69 the vast majority of Mtb epitope-specific T cells in LTBI subjects [11], are associated with
70 increased susceptibility to Mycobacteria infections in humans [12].

71 Here we asked if the transcriptome data of purified memory $CD4^+$ T cells in latent TB infected
72 individuals contained a gene signature associated with risk of disease progression that could help
73 understand the changing role of $CD4^+$ T cells with progression to active disease. We used a
74 validated marker of TB risk progression that was available to us, namely the
75 Monocyte/Lymphocyte (M/L) ratio. Dysregulation of M/L ratio is a well known feature of active
76 TB. It has been described as early as in 1929, where patients with active TB showed high M/L
77 ratio at the onset of disease, that decreased to normal values as healing occurred [13]. More
78 recently, M/L ratio has shown good discriminatory power to identify active TB progressors within
79 LTBI affected adults (HIV seropositive [14] or seronegative [15]), and infants [16]. We found that
80 LTBI subjects with a high M/L ratio (hence higher risk of progression to disease) indeed had a
81 distinct gene expression profile in their $CD4^+$ memory T cells, which was surprisingly highly
82 enriched for monocyte-associated genes. We ultimately traced this discovery to a general
83 phenomenon of circulating T-cell:monocyte complexes in the blood that can be detected following
84 immune perturbations, and could serve as a general diagnostic and mechanistic tool to better
85 monitor and understand T cell:APC immune interactions.

86

87 **Results**

88 Unexpected detection of monocyte gene expression in CD4$^+$ memory T cells from M/L high
89 subjects.

90 To identify individuals at high risk of progression to active disease in our previously characterized
91 cohort of LTBI individuals [6], we determined the frequency of monocytes and lymphocytes in
92 their cryopreserved PBMC using our previously published 10-color flow cytometry panel [17].
93 M/L ratio was calculated based on forward and side scatter identification of the monocyte and
94 lymphocyte populations (**Fig. S2A**), and showed high inter-individual variability within the LTBI
95 cohort (**Fig. S2B**). To identify biological relevant thresholds to apply to the LTBI cohort, we
96 calculated the M/L ratio from a TB negative cohort of 29 individuals [6] in the same fashion. The
97 highest M/L ratio amongst any TB negative subjects was used as a threshold to classify LTBI
98 individuals as M/L high (>0.56, n=6 subjects). The median M/L ratio in the TB negative cohort
99 was then used to classify the remaining LTBI individuals into M/L low (< 0.31, n=12 subjects)
100 and M/L intermediate (M/L int) (0.31 – 0.56, n=12 subjects) (**Fig. 1A**, **Fig. S2B**). Thus, based on
101 the M/L ratio from Mtb uninfected individuals, we could group our LTBI cohort into individuals
102 with presumed high, intermediate, and low risk to develop active TB.

103 We wanted to determine if there was a signature in the previously obtained transcriptomic profiles
104 of memory CD4$^+$ T cells in our LTBI cohort that distinguishes high risk from low risk individuals.
105 Weighted Gene-Coexpression Network Analysis (WGCNA [18]) identified 10 modules
106 containing genes with correlated expression patterns. For each module, we compared the overall
107 level of expression of genes in the module for a given individual with the M/L ratio. This identified
108 one module containing 53 genes that significantly positively correlated with the M/L ratio
109 (spearman correlation = 0.49, p value = 0.006, **Fig. S3A**), and which could readily discriminate

110 M/L high from M/L low individuals (**Fig. 1B**). Surprisingly, many of the genes contained within
111 the 53-gene module were previously described as being highly expressed in monocytes but not in
112 T cells (**Fig. S3B, Table S2**). In particular, the 53-gene module contained the commonly used
113 monocyte lineage marker CD14, the monocyte/macrophage lineage marker CD68, and the
114 granulocytes/monocyte S100 calcium binding protein S100A9. All of these genes showed
115 increased expression levels in M/L high compared to M/L low subjects (**Fig. 1C**). By examining
116 the fcs files that were acquired during cell sorting (**Fig. S1A**), we identified that indeed there was
117 a subpopulation within sorted memory $CD4^+$ T cells that stained positive for CD14 (**Fig. 1D**). The
118 proportion of memory $CD4^+$ T cells that were $CD14^+$ was higher in M/L high compared to M/L
119 low subjects (**Fig. 1D**), suggesting that this cell subset is responsible for the expression of the 53-
120 gene monocyte module identified in Fig. 1B. The $CD14^+$ memory $CD4^+$ T cell population had
121 similar forward and side scatter (FSC/SSC) values to other memory $CD4^+$ T cells and was thus
122 sorted along with conventional $CD14^-$ memory $CD4^+$ T cells (**Fig. S4**). In particular, there was no
123 indication that the $CD14^+$ memory $CD4^+$ T cells were the product of a technical artefact, such as
124 dead cells or a compensation issue.

125 <u>Two distinct $CD3^+CD14^+$ cell populations are present in the monocyte and lymphocyte size gate.</u>
126 When gating on live FSC/SSC (including both monocytes and lymphocytes) cells, two populations
127 of $CD3^+CD14^+$ were identified: $CD3^+CD14^{hi}$ cells and $CD3^+CD14^{mid}$ cells (**Fig. 2A**). $CD3^+CD14^{hi}$
128 cells were predominantly contained within the monocyte size gate (Mgate), whereas $CD3^+CD14^{mid}$
129 cells were contained within the lymphocyte size gate (Lgate) (**Fig. 2B**). Thus, $CD3^+CD14^{hi}$ and
130 $CD3^+CD14^{mid}$ cells can also be identified as '$CD3^+CD14^+$ Mgate' and '$CD3^+CD14^+$ Lgate,
131 respectively. Similar to $CD14^+$ memory $CD4^+$ T cells, the frequency of total live $CD3^+CD14^+$ cells
132 was higher in M/L high compared to M/L low subjects (**Fig. 2C**). However, the relative proportion

133 of $CD3^+CD14^+$ cells in the Lgate and Mgate was independent of M/L ratio (**Fig. 2D**), suggesting
134 that both cell populations are generated through a common mechanism.

135 To examine how the gene expression program in $CD3^+CD14^+$ cells compares to conventional T
136 cells and monocytes, we performed single-cell RNA sequencing. Single cells from monocytes
137 ($CD3^-CD14^+$), T cells ($CD3^+CD14^-$), $CD3^+CD14^+$ Mgate cells and $CD3^+CD14^+$ Lgate cells were
138 sorted from one LTBI subject. Principal Component Analysis (PCA) of the entire set of genes
139 showed that $CD3^+CD14^+$ Mgate cells represent an intermediate population between T cells and
140 monocytes, whereas $CD3^+CD14^+$ Lgate cells represent an intermediate population between T cells
141 and $CD3^+CD14^+$ Mgate cells (**Fig. 2E**). The 53-gene monocyte module previously identified (Fig.
142 1C) was sufficient to distinguish between the four sorted cell populations with gradual expression
143 from high to low within monocytes, $CD3^+CD14^+$ Mgate, $CD3^+CD14^+$ Lgate and T cells,
144 respectively (**Fig. 2F**). This further confirmed that the $CD3^+CD14^+$ Mgate- and Lgate- populations
145 are distinct, and seem to show an intermediate gene expression profile between monocytes and T
146 cells.

147 <u>$CD3^+CD14^+$ cells consist of T cells bound to monocytes or monocyte debris.</u>
148 We aimed to visualize the $CD3^+CD14^+$ cells and the distribution of their markers using imaging
149 flow cytometry. Live events were divided into monocytes ($CD3^-CD14^+$), T cells ($CD3^+CD14^-$),
150 $CD3^+CD14^{hi}$ cells, and $CD3^+CD14^{mid}$ cells (**Fig. 3A**), and a random gallery of images was
151 exported for each population. As expected, monocytes and T cells contained exclusively single
152 cells that expressed either CD14 (monocytes) or CD3 (T cells), respectively (**Fig. 3B**, *first and*
153 *second panel*). To our surprise, $CD3^+CD14^{hi}$ cells contained predominantly two cells, sometimes
154 even three cells, but no single cells (**Fig. 3B**, *third panel*). The doublets (or triplets) always
155 contained at least one $CD14^+$ cell, and one $CD3^+$ cell (**Fig. 3B**, *third panel*). $CD3^+CD14^{mid}$ cells

156 contained predominantly single $CD3^+$ cells, but also some doublets of one $CD3^+$ cell and one
157 $CD14^+$ cell, but with CD14 expression lower than average monocytes, and these doublets were
158 present in the upper part of the $CD14^{mid}$ gate (**Fig. 3B**, *fourth panel*). The majority of $CD3^+$ T cell
159 singlets in the $CD3^+CD14^{mid}$ population, but not in the $CD3^+CD14^-$ T cell population, contained
160 smaller $CD14^+$ particles, often seen at the periphery of the $CD3^+$ T cell membrane (**Fig. 3B** and
161 **3C**). Looking more closely at the $CD14^+$ particles contained within the $CD3^+CD14^{mid}$ population
162 using confocal microscopy, they were found to have size and shape similar to cell debris (**Fig. 3D**).
163 Similarly, single sorted monocytes and T cells contained exclusively one $CD14^+$ cell or one $CD3^+$
164 cell, respectively, whereas single sorted $CD3^+CD14^+$ cells contained one $CD3^+$ cell and one $CD14^+$
165 cell (**Fig. S5**). Taken together, these results demonstrate that $CD3^+CD14^{hi}$ cells are tightly bound
166 T cell:monocyte complexes, in such strong interaction that sample processing and flow cytometry
167 acquisition did not break them apart. This conclusion is further supported by $CD3^+CD14^{hi}$
168 complexes being found in the monocyte size gate, and by their RNA expression profile showing a
169 nearly perfect mix between monocytes and T cells. Conversely, the $CD3^+CD14^{mid}$ population
170 appears to predominantly consist of single $CD3^+$ T cells with attached $CD14^+$ cell debris. This is
171 again supported by these cells falling into the lymphocyte size gate, and the single cell RNA
172 sequencing data showing them clustering closer to T cells (Fig. 2D).

173 <u>T cell:monocyte complexes show active T-cell receptor signaling.</u>

174 Next, we sought to determine whether the association between T cells and monocytes within the
175 T cell:monocyte complexes was the result of random cellular proximity during sample freezing or
176 preparation (and thus a technical artefact), or if the complexes represent a true biological interface.
177 A key interaction between T cells and monocytes is the formation of immune synapses between
178 antigen-presenting monocytes and antigen-specific T cells. Upon formation of immune synapses,

179 T-cell receptor (TCR) signaling is triggered in T cells through the binding of antigen/MHC
180 complex with the TCR. In human T cells, TCRz, Zap70 and LAT are the earliest signaling
181 molecules downstream of TCR activation [19]. Binding of the MHC/peptide complex to TCR
182 triggers the recruitment of these molecules to the TCR and their phosphorylation. TCRz and Zap70
183 were both present at higher frequencies in $CD3^{+}CD14^{hi}$ complexes compared to individual T cells,
184 whereas LAT was expressed at similar levels between both T cell containing subsets (**Fig. 3E**).
185 Importantly, all three molecules presented a higher degree of phosphorylation in $CD3^{+}CD14^{hi}$
186 complexes compared to T cells not present in a complex (**Fig. 3E** and **3F**). Activation was much
187 more pronounced in the $CD3^{+}CD14^{hi}$ population containing the intact Tcell:monocyte pairs
188 compared to $CD3^{+}CD14^{mid}$ population containing T cells and monocyte debris (**Fig. 3E** and **3F**),
189 so we focused on this intact cell pairs population for the remainder of this study.

190 Beside TCR/MHC binding, adhesion molecules are critical for the formation of immune synapses
191 [20]. In particular, upon interaction with ICAM1 on the antigen-presenting cell (APC), the integrin
192 LFA-1 on the T cell surface undergo a drastic redistribution by focusing almost exclusively at the
193 cell:cell point of contact, thus forming a 'ring' that can be visualized [21]. Using imaging flow
194 cytometry, we found a drastic increase in LFA-1 fluorescence intensity in the $CD3+CD14^{hi}$
195 complexes, compared to singlets T cells or singlets monocytes (**Fig. 3G**). Looking within the
196 'LFA-1 bright' $CD3^{+}CD14^{hi}$ doublets, we observed that almost the totality of LFA-1 expression
197 was clustered at one specific location on the T cell, corresponding to the point of contact with the
198 monocyte (**Fig. 3H**). Overall, this suggests that T cell:monocyte complexes are enriched for
199 immune synaptic interactions, and that the T cells in the intact complexes are likely to be in the
200 process of recognizing antigen presented to them by the monocytes.

201 T cell:monocyte complexes are increased post vaccination and in severe dengue infection.

202 Our data suggest circulating T cell:monocyte complexes detected at higher frequency in LTBI
203 subjects with high M/L ratio are biologically meaningful and the result of immune synaptic
204 interactions. Thus, they should also be present in other situations where active antigen presentation
205 by monocytes to T cells is taking place. To determine if this was true, we looked at the changes in
206 frequency of T cell:monocyte complexes in two very different immune perturbation models: First,
207 in tetanus, diphtheria and pertussis (Tdap) booster vaccination in healthy adults from San Diego,
208 we indeed observed a significantly higher frequency of $CD3^+CD14^+$ complexes one day post boost,
209 which resumes to baseline levels at seven days post boost (**Fig. 4A**). Second, we found that in
210 subjects with acute dengue fever from Sri Lanka, patients that developed hemorrhagic fever (the
211 more severe form of dengue) had higher frequency of $CD3^+CD14^+$ upon hospitalization compared
212 to convalescent subjects, as opposed to patients that had a less severe forms of acute dengue
213 infection (**Fig. 4B**). Importantly, the distinction between hemorrhagic vs. non-hemorrhagic fever
214 may become clear days into hospitalization, so the ability to discriminate these two groups of
215 individuals based on $CD3^+CD14^+$ frequency at the time of admission has potential diagnostic
216 value. Taken together, these data confirmed that circulating T cell:monocyte complexes can be
217 found directly *ex vivo* in different immune perturbations, and their frequency is associated with
218 clinical parameters such as time post vaccination and disease severity.

219 The T cell phenotype within T cell:monocyte complexes reflect pathogen-specific polarization of
220 immune responses.

221 The phenotype of T cells involved in antigen recognition for the different immune perturbations
222 studied is expected to be distinct. Tdap vaccine contains exclusively protein antigens and is thus
223 expected to elicit predominantly $CD4^+$ T cell responses [22]. Mtb is a bacterial pathogen known

224 to trigger strong $CD4^+$ responses [23] as opposed to dengue virus, which is a viral antigen. Thus,
225 if the T cell:monocyte complexes are the result of antigen specific immune perturbations, their T
226 cell phenotype should vary based on the nature of the perturbation. We indeed found that in LTBI
227 the majority of complexes contained $CD4^+$ T cells, which was even more pronounced in the Tdap
228 boost model as expected (**Fig. 4C.**). In contrast, in dengue acute individuals, there were much less
229 $CD4^+$ complexes, but interestingly a shift towards $CD4^-CD8^-$ double negative cells (**Fig. 4C**).
230 These cells could constitute gamma-delta T cells that are known to be strongly activated in the
231 peripheral blood during acute dengue fever [24]. Taken together, these data show that the
232 phenotype of T cells contained within $CD3^+CD14^{hi}$ complexes follows differential pattern of CD4
233 vs. CD8 polarization in the three systems studied, further supporting the notion that the formation
234 of this complex is the result of T cell immune recognition triggered by the immune perturbation.

235 <u>T cell:monocyte complexes of acute dengue infected subjects are enriched for infected monocytes.</u>
236 There are growing evidence that monocytes can uptake and process both cell-associated antigen
237 (through direct infection or phagocytosis of infected cells) and extracellular antigen for
238 presentation to T cells, to a similar fashion to dendritic cell subsets (reviewed in [25]). However,
239 the contribution of monocytes in the process of antigen presentation, in particular for pathogens
240 that infect monocytes, is largely overlooked. In the case of dengue, monocytes are known to get
241 infected, therefore we examined if we could detect the presence of virus in the T cell:monocyte
242 complexes of subjects with acute disease. Staining for the presence of virus by flow cytometry, we
243 indeed found that $CD3^+CD14^{hi}$ complexes contained a higher number of dengue particles
244 compared to monocytes not engaged in a complex (**Fig. 4D**). Thus, identifying pathogens in T
245 cell:monocyte complexes can give insights on the cause and targets of ongoing immune responses.

246

247 **Discussion**

248 In this study, we set out to determine if LTBI subjects with different predicted level of risk of
249 developing active TB (based on M/L ratio) showed alterations in the transcriptional profile of their
250 memory $CD4^+$ T cells, which are involved in controlling the disease. This led to the discovery that
251 individuals at higher risk of developing active TB had an increased frequency of T cells circulating
252 in peripheral blood that were tightly associated with monocytes, and less frequently, with
253 monocyte-derived debris. T cells in the intact complexes showed high levels of TCR signaling and
254 enrichment for immune synaptic markers, suggesting that they were in functional interaction, such
255 as recognizing epitopes presented by the monocytes. Expansion of our studies to other immune
256 perturbations namely Tdap boost vaccination and dengue virus infection again showed increases
257 in circulating T cell:monocyte complexes which correlated with time from vaccination, disease
258 severity, and presence of viral pathogen in the monocytes. Overall, this is consistent with the
259 hypothesis that immune perturbations, even if they originally occur in tissues or lymph nodes, can
260 lead to the generation of T cell:monocyte complexes that can be detected in peripheral blood.

261 Importantly, detecting T cell:monocyte (and likely more broadly T cell:APC) complexes can
262 provide information on the immune perturbations based on which the complexes were formed. In
263 disease patient populations, a higher T cell:monocyte complexes frequency seems to be associated
264 with higher risk for developing active disease in the case of TB, and with severe forms of disease
265 in the case of dengue virus infection. In both cases, it will be necessary to run prospective trials to
266 irrefutably demonstrate that T cell:monocyte complexes frequency have predictive power of the
267 course of disease. Similarly, in vaccine trials, the T cell:APC frequency should be examined as an
268 early readout to gauge how well the immune system has responded to the vaccine. Finally, in
269 apparently 'healthy' populations, or those with diffuse symptoms, an unusually high frequency of

270 circulating complexes in an individual could be used as an indicator of a yet to be determined
271 immune perturbation.

272 Beyond detecting abnormal frequency of T cell:APC complexes, characterizing the T cells and
273 APCs in these complexes can provide further insights into the associated immune perturbation.
274 Our data suggest that there are drastic differences in terms of CD4 and CD8 expression in the
275 detected complexes depending on the immune perturbation studied. Additional T cell
276 characteristics, such as the expression of tissue homing markers or TCR expressed can provide
277 further information on the underlying immune perturbation. Studying the APCs can similarly
278 provide additional information: the cell type, its receptor phenotype, and especially the presence
279 of infectious agents in APCs present in a T cell complex can provide a powerful diagnostic tool.
280 Pathogen detection could be done as in this study using dengue-specific antibodies, or more
281 generally by looking for the presence of bacterial or viral RNA or DNA in the APC.

282 Why were T cell:monocyte complexes not detected and excluded in flow cytometry based on
283 gating strategies to avoid doublets? Surprisingly, all usual parameters (pulse Area (A), Height (H)
284 and Width (W) from forward and side scatter) looked identical between T cell:monocyte intact
285 complexes ($CD3^{+}CD14^{hi}$) and singlets T cells or monocytes (**Fig. S6A**). The only parameter that
286 could readily distinguish between intact $CD3^{+}CD14^{hi}$ complexes and single T cells or monocytes
287 was the brightfield area parameter from the imaging flow cytometer (**Fig. S6B**), which is a feature
288 absent in non-imaging flow cytometry. Thus, it seems that gating approaches and parameters
289 available in conventional flow cytometry are not sufficient to discriminate tightly bound cell pairs
290 from individual cells.

291 Given that T cell:monocyte complexes are not excluded by conventional FACS gating strategies,
292 why were they not reported previously? Examining our own past studies, a major reason is that
293 lineage markers for T cells (CD3), B cells (CD19) and monocytes (CD14) are routinely used to
294 remove cells not of interest in a given experiment by adding them to a 'dump channel'. For
295 example, most of our $CD4^+$ T cell studies have CD8, CD19 and CD14, and dead cell markers
296 combined in the same channel [6, 26]. Other groups studying e.g. $CD14^+$ monocytes are likely to
297 add CD3 to their dump gate. This means that complexes of cells that have two conflicting lineage
298 markers such as CD3 and CD14 will often be removed from studies from the start. Additionally,
299 the detection of complexes by flow cytometry is not straightforward. In our hands, we have found
300 that conventional flow analyzers give low frequency of complexes and poor reproducibility in
301 repeat runs. This is opposed to cell sorters, presumably due to differences in their fluidics systems
302 which puts less stress on cells and does not disrupt complexes as much. Both the routine exclusion
303 of cell populations positive for two conflicting lineage markers and the lack of reproducibility in
304 terms of frequencies for such cell populations on different platforms has likely contributed to them
305 not being reported.

306 Moreover, even if a panel allows for the detection of complexes, and there is a stable assay used
307 to show their presence, there is an assumption in the field that detection of complexes is a result
308 of experimental artefacts. For example, we found a report of double positive $CD3^+CD34^+$ cells
309 detected by flow cytometry in human bone marrow, which followed up this finding and found
310 them to be doublets using microscopy imaging. The authors concluded that these complexes are
311 the product of random association and should be ignored [27]. Their conclusion may well be true
312 for their study, but it highlights a dogma in the field of cytometry that pairs of cells have to be
313 artifacts. Another study described $CD3^+CD20^+$ singlets cells observed by flow cytometry as

314 doublets of T cells and B cells, and also concluded them to be a technical artefact, in the sense that
315 these cells are not singlets double expressing CD3 and CD20 [28]. In this case however, authors
316 pointed out that 'Whether the formation of these doublets is an artefact occurring during staining
317 or is a physiologic process remains to be determined' [28]. We ourselves assumed for a long time
318 that we might have an artefact finding, but given the persistent association of T cell:monocyte
319 complexes frequency and phenotype with clinically and physiologically relevant parameters, we
320 came to a new conclusion: cells are meant to interact with other cells. Thus, detecting and
321 characterizing complexes of cells isolated from tissues and bodily fluids, can provide powerful
322 insights into cell:cell communication events that are missed when studying cells as singlets only.
323

324 **Material and methods**

325 Ethics statement

326 Samples from individuals with LTBI were obtained from the University of California, San Diego
327 Antiviral Research Center clinic (AVRC at UCSD, San Diego) and the Universidad Peruana
328 Cayetano Heredia (UPCH, Peru). Dengue samples were obtained from healthy adult blood donors
329 from the National Blood Center (NBC), Ministry of Health, Colombo, Sri Lanka, in an anonymous
330 fashion as previously described [29]. Acute dengue fever samples have been collected at National
331 Institute of Infectious Diseases, Gothatuwa, Angoda, Sri Lanka and the North Colombo Teaching
332 Hospital, Ragama, in Colombo, Sri Lanka. Longitudinal Tdap booster vaccination samples were
333 obtained from healthy adults from San Diego, USA. Ethical approval to carry out this work is
334 maintained through the La Jolla Institute for Allergy and Immunology Institutional Review Board,
335 the Medical Faculty of the University of Colombo (which served as a National Institutes of Health–
336 approved institutional review board for Genetech) and the John's Hopkins School of Public Health
337 Institutional Review Board (RHG holds dual appointment at UPCH and JHU). All clinical
338 investigations have been conducted according to the principles expressed in the Declaration of
339 Helsinki. All participants, except anonymously recruited blood bank donors in Sri Lanka, provided
340 written informed consent prior to participation in the study.

341 Subjects and samples

342 LTBI status was confirmed in subjects by a positive IFN-γ release assay (IGRA) (QuantiFERON-
343 TB Gold In-Tube, Cellestis or T-SPOT.TB, Oxford Immunotec) and the absence of clinical and
344 radiographic signs of active TB. Mtb uninfected control subjects were confirmed as IGRA
345 negative. Acute and convalescent dengue virus fever samples were classified by detection of virus
346 (PCR+) and/or dengue-specific IgM and IgG in the serum. Laboratory parameters such as platelet

347 and leukocyte counts, hematocrit, hemoglobulin, AST, ALT and if applicable an ultrasound
348 examination of the chest and abdomen or an X-ray were used to further diagnose patients with
349 either dengue fever (DF) or dengue hemorrhagic fever (DHF), a more severe form of disease,
350 according to WHO's guidelines. Longitudinal Tdap booster vaccination samples were obtained
351 from individuals originally vaccinated in childhood, and boosted with the DTP vaccine Tdap
352 (Adacel). Blood samples were collected prior, one day and seven days post boost. For all cohorts,
353 PBMC were obtained by density gradient centrifugation (Ficoll-Hypaque, Amersham Biosciences)
354 from leukapheresis or whole blood samples, according to the manufacturer's instructions. Cells
355 were resuspended to 10 to 50 million cells per mL in FBS (Gemini Bio-Products) containing 10%
356 dimethyl sulfoxide (Sigma) and cryopreserved in liquid nitrogen.

357 Flow cytometry

358 Surface staining of unfixed PBMC was performed as previously described in [17]. Briefly, cells
359 were stained with fixable viability dye eFluor506 (eBiosciences) and various combinations of the
360 antibodies listed in **Table S1** for 20min at room temperature. For detection of dengue virus positive
361 cells, staining was performed as above, using the 2H2 antibody, an IgG2a reactive for the prM/M
362 protein of dengue virus, serotypes 1–4 (IgG2a anti-DENV1-4 prM) [30]. This antibody was
363 obtained from Sujan Shresta (La Jolla Institute, La Jolla, CA, USA) and conjugated to FITC using
364 the Pierce FITC antibody labeling kit (ThermoFisher). For intracellular staining of
365 unphosphorylated and phosphorylated TCR signaling molecules, surface stained PBMC were
366 fixed with 100 µl paraformaldehyde (PFA) 4% for 15min at room temperature, washed twice in
367 PBS, and permeabilized with pre-chilled True-Phos Perm buffer (Biolegend) for 45min at -20°C.
368 After two washes in PBS, cells were stained for intracellular TCR signaling antibodies (**Table S1**)
369 for 30min at room temperature. After the final staining step, cells were washed twice in PBS, and

370 resuspended in FACS buffer (PBS containing 0.5% FBS, 2mM EDTA, pH 8.0) and stored at 4°C
371 protected from light until acquisition. Acquisition was performed on a BD LSR-I cell analyzer
372 (BD Biosciences) or on a BD FACSAria III cell sorter (BD Biosciences). Compensation was
373 realized with single-stained beads (UltraComp eBeads, eBiosciences) in PBS using the same
374 antibody dilution as for the cell staining.

375 Cell sorting

376 PBMC were stained with fixable viability dye eFluor506 (eBiosciences) and various combinations
377 of the antibodies listed in **Table S1** as described in the flow cytometry section above. Cell sorting
378 was performed on a BD Aria III/Fusion cell sorter (BD Biosciences). 100,000 memory $CD4^+$ T
379 cells were sorted into TRIzol LS reagent (Invitrogen) for RNA extraction. Single cell sorts of $CD3^-$
380 $CD14^+$ monocytes, $CD3^+CD14^-$ T cells and $CD3^+CD14^+$ Mgate and Lgate cells for RNA
381 sequencing were performed in a 96-well PCR plate containing 4 µl cell lysis buffer. All gating
382 strategies are available in **Fig. S1**.

383 Imaging flow cytometry

384 For the visualization of $CD3^+CD14^+$ cells, PBMC were stained with CD3-AF488 and CD14-PE,
385 or CD3-AF488, CD14-BV421 and LFA1-PE (see Table S1 for antibody details) as described in
386 the flow cytometry section above. After two washes in PBS, cells were resuspended to $10x10^6$
387 cells/mL in FACS buffer containing 5µg/mL Hoechst (Invitrogen) and 1µg/mL 7-AAD
388 (Biolegend) and stored at 4°C protected from light until acquisition. Acquisition was performed
389 with ImageStreamX MkII (Amnis) and INSPIRE software version 200.1.620.0 at 40X
390 magnification and the lowest speed setting. A minimum of 4,000 $CD3^+CD14^+$ events in focus were
391 collected. Data analysis was performed using IDEAS version 6.2.183.0.

392 Microscopy

393 $CD3^-CD14^+$ monocytes, $CD3^+CD14^-$ T cells and $CD3^+CD14^+$ cells were single- cell sorted as
394 described in the cell sorting section on to a microscopy glass slide. Crystals containing sorting
395 buffer and cells were then visualized on a confocal microscope (ZEISS LSM 880) with 20x 0.8NA
396 objective using transmitted light photon multiplier tube (PMT). Fluorescence signals
397 corresponding to CD3 and CD14 were detected with spectrally tuned PMTs using 488 nm and 594
398 nm excitation lines. Additionally, freshly sorted cells were immobilized using CyGel Sustain
399 (Abcam) according to manufacturer recommendations. Three dimensional rendering of cellular
400 fragments (Fig. 3D.) was created in Imaris 9.1 software (Bitplane).

401 Bulk RNA sequencing and analysis

402 RNA sequencing and analysis of memory $CD4^+$ T cells from LTBI infected subjects was
403 performed as described in [31, 32] and quantified by qPCR as described previously [33]. 5 ng of
404 purified total RNA was used for poly(A) mRNA selection, full length reverse-transcription and
405 amplified for 17 cycles, following the smart-seq2 protocol [31, 32]. After purification with
406 Ampure XP beads (Ratio 0.8:1, Beckmann Coulter) and quantification (Picogreen assay,
407 Invitrogen), 1ng of cDNA was used to prepare a Nextera XT sequencing library with the Nextera
408 XT DNA library preparation and index kits (Illumina). Samples were pooled and sequenced using
409 the HiSeq2500 (Illumina) to obtain at least 12 million 50-bp single-end reads per library. The
410 single-end reads that passed Illumina filters were filtered for reads aligning to tRNA, rRNA, and
411 Illumina adapter sequences. The reads were then aligned to UCSC hg19 reference genome using
412 TopHat (v 1.4.1) [34], filtered for low complexity reads, and parsed with SAMtools [35]. Read
413 counts to each genomic feature were obtained using HTSeq-count program (v 0.6.0) [36] using the
414 "union" option. Raw counts were then imported to R/Bioconductor package DESeq2 [37] to

415 identify differentially expressed genes among samples. Modular analysis was performed using
416 automatic network construction and module detection implemented in WGCNA R package [18]
417 following the steps described in package authors' guidelines.

418 Single-cell RNA sequencing and analysis

419 Single cell RNA sequencing was performed using Smart-seq2 as previously described [31]. Single
420 cells were sorted by flow cytometry into 96-well PCR plates containing 4 µl cell lysis buffer. RNA
421 loss was minimized by performing on-plate RNA capture, reverse-transcription and whole
422 transcriptome pre-amplification (24 cycles) that results in ~1–30 ng of cDNA as previously
423 described [31]. 0.3 to 0.5 ng of pre-amplified cDNA was used to generate barcoded Illumina
424 sequencing libraries (Nextera XT library preparation kit, Illumina) in an 8µL reaction volume.
425 Multiple quality-control steps were included to ensure the consistency during the procedure for all
426 samples. Samples that failed at quality-control steps as described in [31, 38], were eliminated from
427 downstream procedures and analysis. Libraries were pooled and sequenced on the HiSeq2500
428 Illumina platform to obtain more than 2 million 50-bp single end reads per cell. Single-cell RNA-
429 seq data were mapped against the human hg19 reference genome using TopHat (v1.4.1., - -library-
430 type fr-secondstrand -C) and Gencode version 19 (GRCh37.p13) as gene model reference for
431 alignment. Sequencing read coverage per gene was counted using HTSeq-count (-m union -s yes
432 -t exon -i gene_id, http://www huber.embl.de/users/anders/HTSeq/). Counts per gene are obtained
433 by counting all the transcripts mapping to a gene and is together referred to as transcript. To
434 normalize for sequencing depth and varying transcript length counts were TPM (Transcripts Per
435 Million reads) transformed. For PCA analysis only genes with at least 20 TPM were considered.
436 PCA was performed in Python using Scipy SVD algorithm and plotted with Matplotlib [39].

437

438 **References**

439 1. The World Health Organization: Global tuberculosis report 2017. Available from:
440 http://www.who.int/tb/publications/global_report/en/.

441 2. Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC, et al. Tuberculosis. Nat
442 Rev Dis Primers. 2016;2:16076. Epub 2016/10/28. doi: 10.1038/nrdp.2016.76. PubMed PMID:
443 27784885.

444 3. Seshadri P, Denkinger C. Draft Target Product Profile: Test for progression of tuberculosis
445 infection 2016. Available from: http://www.finddx.org/wp-content/uploads/2016/05/TPP-
446 LTBIprogression.pdf].

447 4. Zak DE, Penn-Nicholson A, Scriba TJ, Thompson E, Suliman S, Amon LM, et al. A blood
448 RNA signature for tuberculosis disease risk: a prospective cohort study. Lancet.
449 2016;387(10035):2312-22. Epub 2016/03/28. doi: 10.1016/s0140-6736(15)01316-1. PubMed
450 PMID: 27017310; PubMed Central PMCID: PMCPMC5392204.

451 5. Suliman S, Thompson E, Sutherland J, Weiner Rd J, Ota MOC, Shankar S, et al. Four-gene
452 Pan-African Blood Signature Predicts Progression to Tuberculosis. Am J Respir Crit Care Med.
453 2018. Epub 2018/04/07. doi: 10.1164/rccm.201711-2340OC. PubMed PMID: 29624071.

454 6. Burel JG, Lindestam Arlehamn CS, Khan N, Seumois G, Greenbaum JA, Taplitz R, et al.
455 Transcriptomic Analysis of CD4(+) T Cells Reveals Novel Immune Signatures of Latent
456 Tuberculosis. J Immunol. 2018;200(9):3283-90. Epub 2018/04/01. doi:
457 10.4049/jimmunol.1800118. PubMed PMID: 29602771.

458 7. Burel JG, Peters B. Discovering transcriptional signatures of disease for diagnosis versus
459 mechanism. Nat Rev Immunol. 2018;18(5):289-90. Epub 2018/04/10. doi: 10.1038/nri.2018.26.
460 PubMed PMID: 29629714.

461 8. Barnes PF, Bloch AB, Davidson PT, Snider DE, Jr. Tuberculosis in patients with human
462 immunodeficiency virus infection. N Engl J Med. 1991;324(23):1644-50. Epub 1991/06/06. doi:
463 10.1056/nejm199106063242307. PubMed PMID: 2030721.

464 9. Flynn JL, Chan J. Immunology of tuberculosis. Annu Rev Immunol. 2001;19:93-129. Epub
465 2001/03/13. doi: 10.1146/annurev.immunol.19.1.93. PubMed PMID: 11244032.

466 10. Fletcher HA, Snowden MA, Landry B, Rida W, Satti I, Harris SA, et al. T-cell activation
467 is an immune correlate of risk in BCG vaccinated infants. Nat Commun. 2016;7:11290. Epub
468 2016/04/14. doi: 10.1038/ncomms11290. PubMed PMID: 27068708; PubMed Central PMCID:
469 PMCPMC4832066.

470 11. Lindestam Arlehamn CS, Gerasimova A, Mele F, Henderson R, Swann J, Greenbaum JA,
471 et al. Memory T cells in latent Mycobacterium tuberculosis infection are directed against three

472 antigenic islands and largely contained in a CXCR3+CCR6+ Th1 subset. PLoS Pathog.
473 2013;9(1):e1003130. Epub 2013/01/30. doi: 10.1371/journal.ppat.1003130. PubMed PMID:
474 23358848; PubMed Central PMCID: PMCPMC3554618.

475 12. Okada S, Markle JG, Deenick EK, Mele F, Averbuch D, Lagos M, et al.
476 IMMUNODEFICIENCIES. Impairment of immunity to Candida and Mycobacterium in humans
477 with bi-allelic RORC mutations. Science. 2015;349(6248):606-13. Epub 2015/07/15. doi:
478 10.1126/science.aaa4282. PubMed PMID: 26160376; PubMed Central PMCID:
479 PMCPMC4668938.

480 13. Blackfan KDD, L.K. The monocyte in active tuberculosis - supravital studies of the blood.
481 Am J Dis Child. 1929;37(2):233-43. doi: doi:10.1001/archpedi.1929.01930020003001.

482 14. Naranbhai V, Hill AV, Abdool Karim SS, Naidoo K, Abdool Karim Q, Warimwe GM, et
483 al. Ratio of monocytes to lymphocytes in peripheral blood identifies adults at risk of incident
484 tuberculosis among HIV-infected adults initiating antiretroviral therapy. J Infect Dis.
485 2014;209(4):500-9. Epub 2013/09/18. doi: 10.1093/infdis/jit494. PubMed PMID: 24041796;
486 PubMed Central PMCID: PMCPMC3903371.

487 15. Rakotosamimanana N, Richard V, Raharimanga V, Gicquel B, Doherty TM, Zumla A, et
488 al. Biomarkers for risk of developing active tuberculosis in contacts of TB patients: a prospective
489 cohort study. Eur Respir J. 2015;46(4):1095-103. Epub 2015/08/08. doi:
490 10.1183/13993003.00263-2015. PubMed PMID: 26250497.

491 16. Naranbhai V, Kim S, Fletcher H, Cotton MF, Violari A, Mitchell C, et al. The association
492 between the ratio of monocytes:lymphocytes at age 3 months and risk of tuberculosis (TB) in the
493 first two years of life. BMC Med. 2014;12:120. Epub 2014/07/19. doi: 10.1186/s12916-014-0120-
494 7. PubMed PMID: 25034889; PubMed Central PMCID: PMCPMC4223414.

495 17. Burel JG, Qian Y, Lindestam Arlehamn C, Weiskopf D, Zapardiel-Gonzalo J, Taplitz R,
496 et al. An Integrated Workflow To Assess Technical and Biological Variability of Cell Population
497 Frequencies in Human Peripheral Blood by Flow Cytometry. J Immunol. 2017;198(4):1748-58.
498 Epub 2017/01/11. doi: 10.4049/jimmunol.1601750. PubMed PMID: 28069807; PubMed Central
499 PMCID: PMCPMC5296239.

500 18. Zhang B, Horvath S. A general framework for weighted gene co-expression network
501 analysis. Stat Appl Genet Mol Biol. 2005;4:Article17. Epub 2006/05/02. doi: 10.2202/1544-
502 6115.1128. PubMed PMID: 16646834.

503 19. Chakraborty AK, Weiss A. Insights into the initiation of TCR signaling. Nat Immunol.
504 2014;15(9):798-807. Epub 2014/08/20. doi: 10.1038/ni.2940. PubMed PMID: 25137454; PubMed
505 Central PMCID: PMCPMC5226627.

506 20. Dustin ML. The immunological synapse. Cancer Immunol Res. 2014;2(11):1023-33. Epub
507 2014/11/05. doi: 10.1158/2326-6066.cir-14-0161. PubMed PMID: 25367977; PubMed Central
508 PMCID: PMCPMC4692051.

509 21. Wabnitz GH, Samstag Y. Multiparametric Characterization of Human T-Cell Immune
510 Synapses by InFlow Microscopy. Methods Mol Biol. 2016;1389:155-66. Epub 2016/07/28. doi:
511 10.1007/978-1-4939-3302-0_10. PubMed PMID: 27460243.

512 22. Bancroft T, Dillon MB, da Silva Antunes R, Paul S, Peters B, Crotty S, et al. Th1 versus
513 Th2 T cell polarization by whole-cell and acellular childhood pertussis vaccines persists upon re-
514 immunization in adolescence and adulthood. Cell Immunol. 2016;304-305:35-43. Epub
515 2016/05/24. doi: 10.1016/j.cellimm.2016.05.002. PubMed PMID: 27212461; PubMed Central
516 PMCID: PMCPMC4899275.

517 23. Lindestam Arlehamn CS, McKinney DM, Carpenter C, Paul S, Rozot V, Makgotlho E, et
518 al. A Quantitative Analysis of Complexity of Human Pathogen-Specific CD4 T Cell Responses in
519 Healthy M. tuberculosis Infected South Africans. PLoS Pathog. 2016;12(7):e1005760. Epub
520 2016/07/15. doi: 10.1371/journal.ppat.1005760. PubMed PMID: 27409590; PubMed Central
521 PMCID: PMCPMC4943605.

522 24. Tsai CY, Liong KH, Gunalan MG, Li N, Lim DS, Fisher DA, et al. Type I IFNs and IL-18
523 regulate the antiviral response of primary human gammadelta T cells against dendritic cells
524 infected with Dengue virus. J Immunol. 2015;194(8):3890-900. Epub 2015/03/04. doi:
525 10.4049/jimmunol.1303343. PubMed PMID: 25732728.

526 25. Jakubzick CV, Randolph GJ, Henson PM. Monocyte differentiation and antigen-presenting
527 functions. Nat Rev Immunol. 2017;17(6):349-62. Epub 2017/04/25. doi: 10.1038/nri.2017.28.
528 PubMed PMID: 28436425.

529 26. Arlehamn CL, Seumois G, Gerasimova A, Huang C, Fu Z, Yue X, et al. Transcriptional
530 profile of tuberculosis antigen-specific T cells reveals novel multifunctional features. J Immunol.
531 2014;193(6):2931-40. Epub 2014/08/06. doi: 10.4049/jimmunol.1401151. PubMed PMID:
532 25092889; PubMed Central PMCID: PMCPMC4157075.

533 27. Kudernatsch RF, Letsch A, Stachelscheid H, Volk HD, Scheibenbogen C. Doublets
534 pretending to be CD34+ T cells despite doublet exclusion. Cytometry A. 2013;83(2):173-6. Epub
535 2013/01/03. doi: 10.1002/cyto.a.22247. PubMed PMID: 23281028.

536 28. Henry C, Ramadan A, Montcuquet N, Pallandre JR, Mercier-Letondal P, Deschamps M,
537 et al. CD3+CD20+ cells may be an artifact of flow cytometry: comment on the article by Wilk et
538 al. Arthritis Rheum. 2010;62(8):2561-3; author reply 3-5. Epub 2010/05/28. doi:
539 10.1002/art.27527. PubMed PMID: 20506165.

540 29. Weiskopf D, Angelo MA, de Azeredo EL, Sidney J, Greenbaum JA, Fernando AN, et al.
541 Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective
542 role for CD8+ T cells. Proc Natl Acad Sci U S A. 2013;110(22):E2046-53. Epub 2013/04/13. doi:
543 10.1073/pnas.1305227110. PubMed PMID: 23580623; PubMed Central PMCID:
544 PMCPMC3670335.

30. Zellweger RM, Miller R, Eddy WE, White LJ, Johnston RE, Shresta S. Role of humoral versus cellular responses induced by a protective dengue vaccine candidate. PLoS Pathog. 2013;9(10):e1003723. Epub 2013/11/10. doi: 10.1371/journal.ppat.1003723. PubMed PMID: 24204271; PubMed Central PMCID: PMCPMC3814346.

31. 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(11):1096-8. Epub 2013/09/24. doi: 10.1038/nmeth.2639. PubMed PMID: 24056875.

32. Seumois G, Zapardiel-Gonzalo J, White B, Singh D, Schulten V, Dillon M, et al. Transcriptional Profiling of Th2 Cells Identifies Pathogenic Features Associated with Asthma. J Immunol. 2016;197(2):655-64. Epub 2016/06/09. doi: 10.4049/jimmunol.1600397. PubMed PMID: 27271570; PubMed Central PMCID: PMCPMC4936908.

33. Seumois G, Vijayanand P, Eisley CJ, Omran N, Kalinke L, North M, et al. An integrated nano-scale approach to profile miRNAs in limited clinical samples. Am J Clin Exp Immunol. 2012;1(2):70-89. Epub 2013/01/11. PubMed PMID: 23304658; PubMed Central PMCID: PMCPMC3538381.

34. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 2009;25(9):1105-11. Epub 2009/03/18. doi: 10.1093/bioinformatics/btp120. PubMed PMID: 19289445; PubMed Central PMCID: PMCPMC2672628.

35. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9. Epub 2009/06/10. doi: 10.1093/bioinformatics/btp352. PubMed PMID: 19505943; PubMed Central PMCID: PMCPMC2723002.

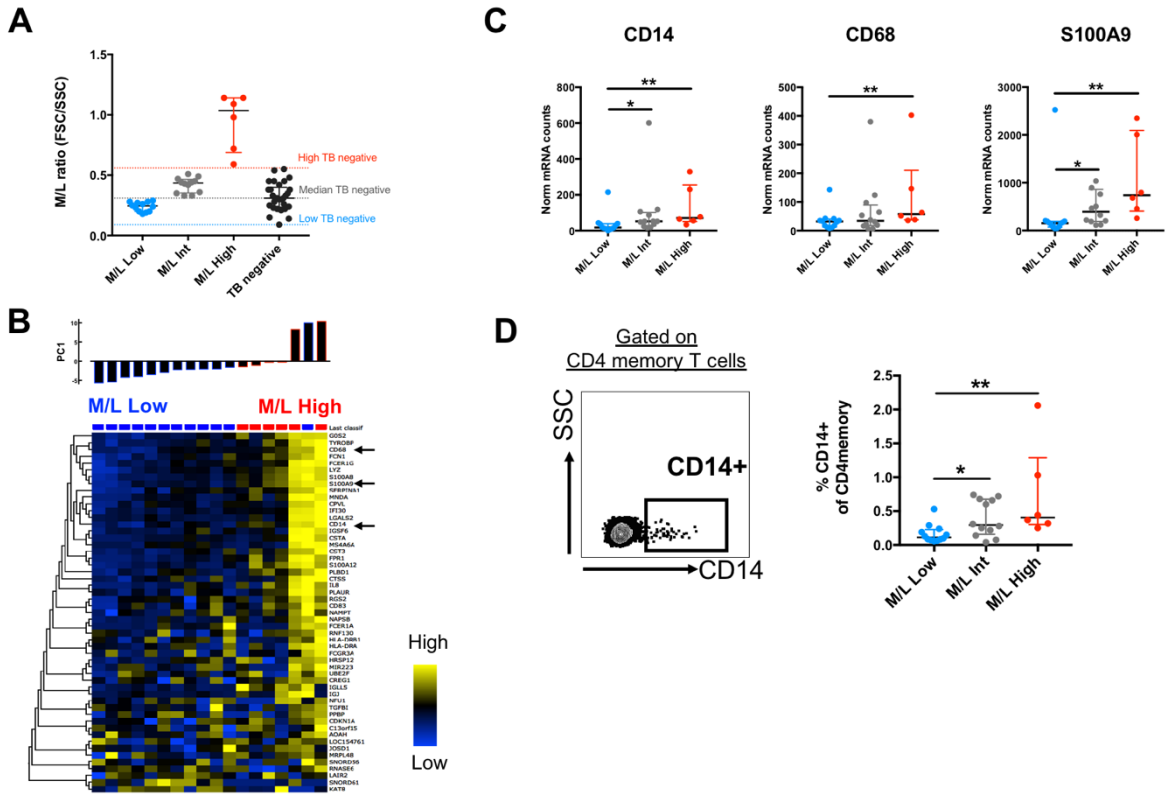
36. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166-9. Epub 2014/09/28. doi: 10.1093/bioinformatics/btu638. PubMed PMID: 25260700; PubMed Central PMCID: PMCPMC4287950.

37. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. Epub 2014/12/18. doi: 10.1186/s13059-014-0550-8. PubMed PMID: 25516281; PubMed Central PMCID: PMCPMC4302049.

38. Patil VS, Madrigal A, Schmiedel BJ, Clarke J, O'Rourke P, de Silva AD, et al. Precursors of human CD4(+) cytotoxic T lymphocytes identified by single-cell transcriptome analysis. Sci Immunol. 2018;3(19). Epub 2018/01/21. doi: 10.1126/sciimmunol.aan8664. PubMed PMID: 29352091; PubMed Central PMCID: PMCPMC5931334.

39. Hunter JD. Matplotlib: a 2D Graphics Environment. Computing in Science & Engineering. 2007;9(3):90-5. doi: 10.1109/MCSE.2007.55.

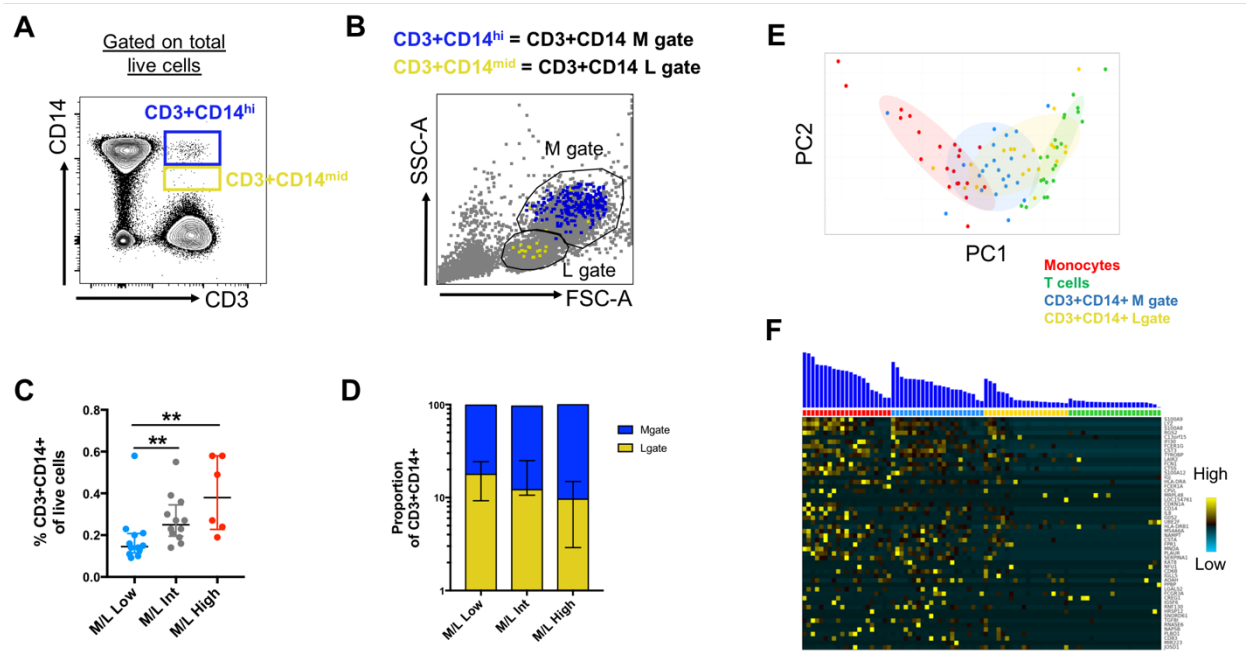
582 **Acknowledgments:** We thank the flow cytometry core facility at the La Jolla Institute for
583 assistance in cell sorting and technical discussion (in particular Dr. Cheryl Kim and Dr. Denise
584 Hinz). We thank Dr. Zbigniew Mikulski from the microscopy core at the La Jolla Institute for
585 assistance and technical advice on microscopy imaging. We thank Yoav Altman at the Sanford
586 Burnham Prebys flow cytometry core for technical assistance with imaging flow cytometry. We
587 thank Dr. Joe Trotter from the R&D Advanced Technology Group at BD Biosciences for useful
588 technical discussion about cytometry instrument fluidic systems. We thank Dr. Klaus Ley from
589 the La Jolla Institute for multiple suggestions on experimental design. We thank Dr. Amnon
590 Altman from the La Jolla Institute for TCR signaling experimental design. **Funding:** Research
591 reported in this manuscript was supported by the National Institute of Allergy and Infectious
592 Diseases division of the National Institutes of Health under award number U19AI118626,
593 R01AI137681, HHSN272201400045C, S10OD021831 and S10OD016262. The content is solely
594 the responsibility of the authors and does not necessarily represent the official views of the
595 National Institutes of Health. Imaging flow cytometry was supported by the James B. Pendleton
596 Charitable trust. **Author contributions:** conceptualization: JGB, MP, CLA, MB, AS and BP;
597 investigation: JGB, MP, CLA, DW, RSA, VS and GS; resources: CLA, JAG, SP, GP, AW, RT,
598 ADS, RH, MS, RT and PV; writing (original draft preparation): JGB and BP; writing (review and
599 editing): all authors; funding acquisition: AS and BP. **Competing interests:** Authors declare no
600 competing interests. **Data and materials availability:** Sequencing data is accessible online
601 through Gene Expression Omnibus (accession numbers GSE84445, GSE99373 and GSE117435,
602 https://www.ncbi.nlm.nih.gov/geo) and Immport (Study number SDY820,
603 http://www.immport.org). All other data is available in the main text or the supplementary
604 materials.



605

606 **Fig. 1. Unexpected detection of CD14 in sorted CD4$^+$ memory T cells from M/L high LTBI**
607 **subjects.** A) LTBI subjects were stratified based on their Monocyte/Lymphocyte (M/L) ratio
608 measured as the frequency of CD14$^+$ cells from live cells by flow cytometry. B) M/L ratio in LTBI
609 subjects positively correlated with the expression of a 53-gene module in *ex vivo* sorted memory
610 CD4$^+$ T cells. C) CD14, CD68 and S100A9 gene expression in memory CD4$^+$ T cells of M/L high,
611 int and low LTBI subjects. D) The frequency of CD14$^+$ cells within sorted CD4$^+$ memory T cells
612 in M/L high, int and low LTBI subjects. Data were derived from 30 LTBI subjects (6 M/L high,
613 16 M/L int and 8 M/L low) and 29 Mtb negative control subjects. Statistical differences were
614 determined using the non-parametric Wilcoxon test; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

615



616

617 **Fig. 2. Two separate CD3$^+$CD14$^+$ cell populations exist in Mgate and Lgate that form**
618 **intermediate populations between T cells and monocytes.** Cryopreserved PBMC from LTBI
619 subjects were stained for CD3 and CD14 protein expression and acquisition performed on BD Aria
620 III cell sorter. A) Gated on 'singlets' total live cells, two populations of CD3$^+$CD14$^+$ cells can be
621 identified based on the level of expression of CD14. B) Based on FSC and SSC parameters,
622 CD3$^+$CD14^{hi} cells are contained within the monocyte gate (Mgate), whereas CD3$^+$CD14^{mid} cells
623 are contained within the lymphocyte gate (Lgate). C) The frequency of total CD3$^+$CD14$^+$ cells and
624 D) their repartition into Mgate vs Lgate in M/L High (n=6), M/L Int (n=16) and M/L Low (n=8)
625 LTBI subjects. E) PCA on all genes and F) expression of the 53-gene module associated with M/L
626 ratio (determined in Fig. 1B) from RNA single-cell sequencing of CD14$^+$CD3$^-$ monocytes (n=21),
627 CD14$^-$CD3$^+$ T cells (n=22), CD3$^+$CD14$^+$ Mgate cells (n=22) and CD3$^+$CD14$^+$ Lgate cells (n=20)
628 derived from one LTBI subject.

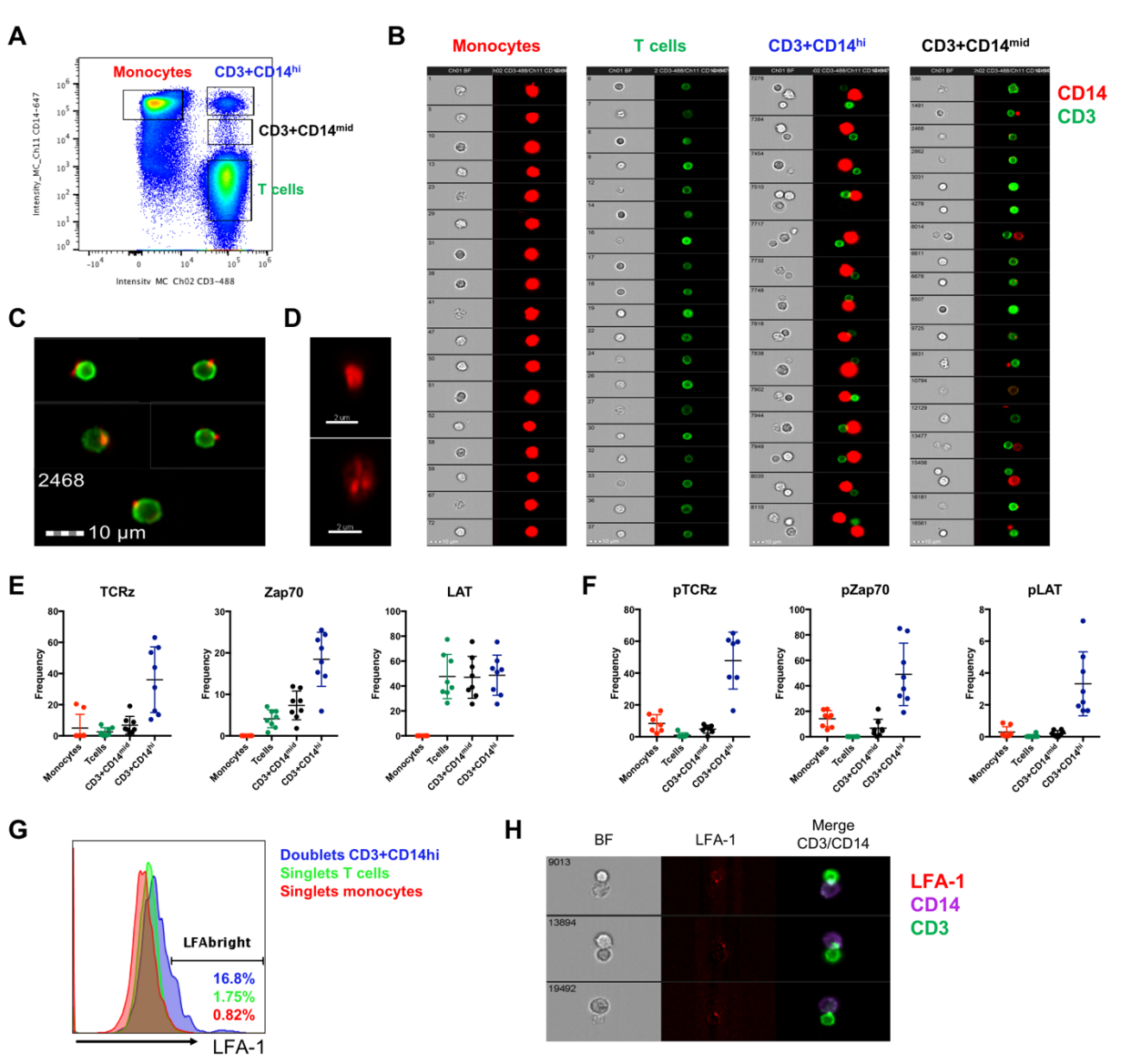


Fig. 3. CD3^{+}CD14^{+} cells are tightly bound T cell:monocyte complexes that do not result from random interaction. A) Gating strategy and B) random gallery of events for monocytes (CD14^{+}CD3^{-}), T cells (CD3^{+}CD14^{-}), CD3^{+}CD14^{hi} cells and CD3^{+}CD14^{mid} cells determined by imaging flow cytometry (ImageStream, Amnis) from one LTBI subject. CD14^{+} cell debris were identified within C) CD3^{+}CD14^{mid} cells by imaging flow cytometry (ImageStreamX MkII, Amnis) and D) CD3^{+}CD14^{+} Lgate cells by confocal microscopy after bulk population cell sorting. Frequency of E) total and F) phosphorylated TCRz, Zap70 and LAT positive cells in monocytes

637 ($CD14^+CD3^-$), T cells ($CD3^+CD14^-$) and $CD3^+CD14^+$ cells determined by phosphoflow.
638 Acquisition was performed on a BD Aria III cell sorter. Data was derived from 8 LTBI subjects
639 and two independent experiments. G) LFA-1 expression (raw pixel intensity) of live doublets
640 $CD3^+CD14^{hi}$, live singlets T cells ($CD3^+CD14^-$) and live singlets monocytes ($CD14^+CD3^-$) and H)
641 representative gallery of live doublets $CD3^+CD14^{hi}$ LFA-1 bright events determined by imaging
642 flow cytometry (ImageStreamX MkII, Amnis) from one LTBI subject. BF: brightfield.

643

644

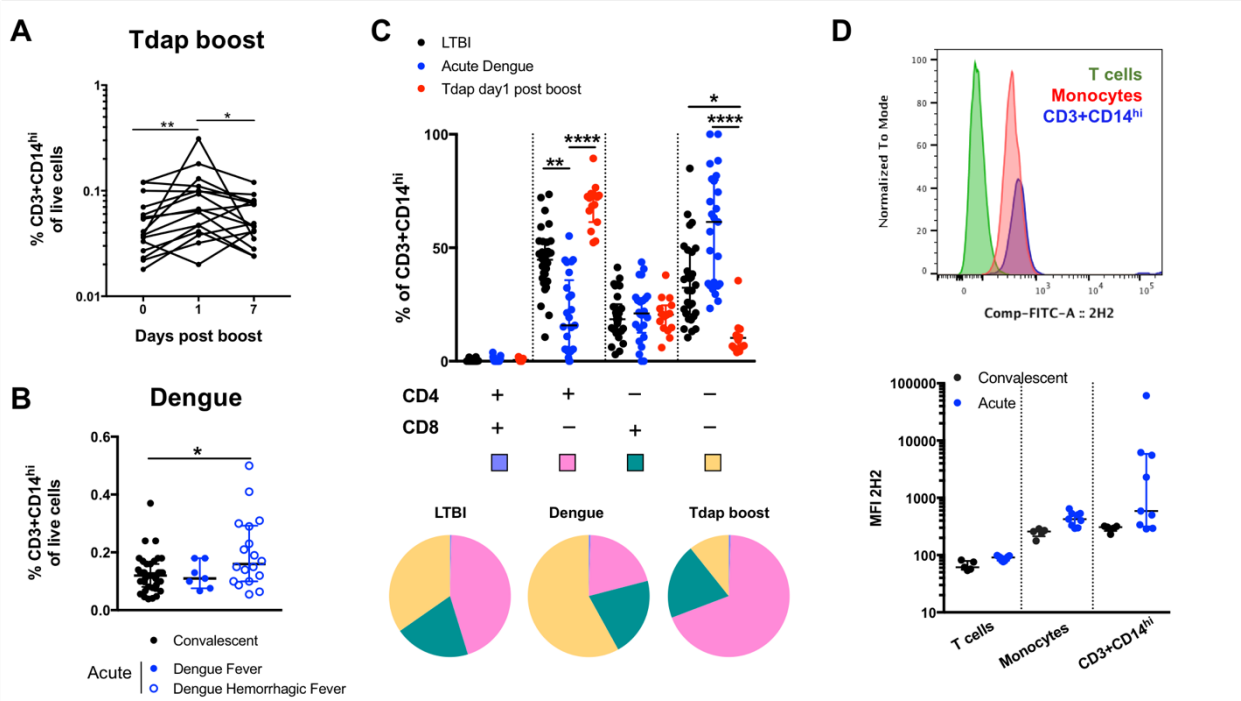


Fig. 4. CD3$^+$CD14$^+$ complexes are also present in other immune perturbations and contain diagnostic and mechanistic information. Frequency of total CD3$^+$CD14$^+$ cells in A) previously vaccinated healthy adults (n=15) before, one day and seven days post-boost with Tdap vaccine and in B) individuals with acute dengue fever (n=7), acute dengue hemorrhagic fever (n=18) or previously infected (convalescent, n=39). C) CD4$^+$ and CD8$^+$ Boolean gating in CD3$^+$CD14$^+$ cells of LTBI subjects (n=30), previously vaccinated healthy adults one day post boost with Tdap vaccine (n=15) and acute dengue virus infected subjects (n=25). D) Dengue virus expression in monocytes (CD14$^+$CD3$^-$), T cells (CD3$^+$CD14$^-$) and CD3$^+$CD14$^+$ cells of subjects with PCR positive acute dengue infection (n=10) and previously infected convalescent subjects (n=5) determined by flow cytometry. Data was acquired with a BD LSR-II (A) or a BD Aria III cell sorter (B-E). In Tdap boosted individuals, statistical differences over time were determined using the non-parametric paired Wilcoxon test; other statistical differences were determined using the non-parametric Wilcoxon test; *, $p < 0.05$; **, $p < 0.01$; ****, $p < 0.0001$.