Multidimensional analysis of Gammaherpesvirus RNA expression reveals unexpected heterogeneity of gene expression

Lauren M. Oko[1], Abigail K. Kimball[2], Rachael E. Kaspar[2], Ashley N. Knox[1], Tim Chang[3], Benjamin Alderete[3], Linda F. van Dyk[1]*, Eric T. Clambey[2]*#

Affiliations:

[1] Department of Immunology and Microbiology, [2] Department of Anesthesiology, University of Colorado Denver | Anschutz Medical Campus, Aurora, CO, 80045, USA

[3] MilliporeSigma, a business of Merck KGaA, Darmstadt, Germany (Seattle, WA)

*** Co-Corresponding authors:**

Eric T. Clambey, eric.clambey@ucdenver.edu, 303-724-7783 (phone)

Linda F. van Dyk, linda.vandyk@ucdenver.edu, 303-724-4207 (phone)

 # Lead author for MS correspondence

Short title: Single-cell heterogeneity of Gammaherpesvirus RNA expression

ABSTRACT

Virus-host interactions are frequently studied in bulk cell populations, obscuring cell-to-cell variation. Here we investigate endogenous herpesvirus gene expression at the single-cell level, combining a sensitive and robust fluorescent in situ hybridization platform with multiparameter flow cytometry, to study the expression of gammaherpesvirus non-coding RNAs (ncRNAs) during lytic replication, latent infection and reactivation in vitro. This method allowed robust detection of viral ncRNAs of murine gammaherpesvirus 68 (γHV68), Kaposi's sarcoma associated herpesvirus and Epstein-Barr virus, revealing variable expression at the single-cell level. By quantifying the inter-relationship of viral ncRNA, viral mRNA, viral protein and host mRNA regulation during γHV68 infection, we find heterogeneous and asynchronous gene expression during latency and reactivation, with reactivation from latency identified by a distinct gene expression profile within rare cells. Further, during lytic replication with γHV68, we find many cells have limited viral gene expression, with only a fraction of cells showing robust gene expression, dynamic RNA localization, and progressive infection. These findings, powered by single-cell analysis integrated with automated clustering algorithms, suggest inefficient or abortive γHV infection in many cells, and identify substantial heterogeneity in viral gene expression at the single-cell level.

AUTHOR SUMMARY

The gammaherpesviruses are a group of DNA tumor viruses that establish lifelong infection. How these viruses infect and manipulate cells has frequently been studied in bulk populations of cells. While these studies have been incredibly insightful, there is limited understanding of how virus infection proceeds within a single cell. Here we present a new approach to quantify gammaherpesvirus gene expression at the single-cell level. This method allows us to detect cell-to-cell variation in the expression of virus non-coding RNAs, an important and understudied class of RNAs which do not encode for proteins. By examining multiple features of virus gene expression, this method further reveals significant variation in infection between cells across multiple stages of infection. These studies emphasize that gammaherpesvirus infection can be surprisingly heterogeneous when viewed at the level of the individual cell. Because this approach can be broadly applied across diverse viruses, this study affords new opportunities to understand the complexity of virus infection within single cells.

INTRODUCTION

The *Herpesviridae* are a family of large dsDNA viruses that include multiple prominent human and animal pathogens [1]. Although these viruses infect different cell types, and are associated with diverse pathologies, they share conserved genes and two fundamental phases of infection: lytic replication and latent infection [1]. Lytic replication is characterized by a cascade of viral gene expression, active viral DNA replication and the production of infectious virions. Conversely, latency is characterized by limited viral gene expression and the absence of de novo viral replication. While

64 latent infection is a relatively quiescent form of infection, the herpesviruses can
65 reactivate from latency, to reinitiate lytic replication.
66 Among the herpesviruses, the gammaherpesviruses (γHV) are lymphotropic
67 viruses that include the human pathogens Epstein-Barr virus (EBV) [2] and Kaposi's
68 sarcoma associated herpesvirus (KSHV) [3]. Murine gammaherpesvirus 68 (γHV68, or
69 MHV-68; ICTV nomenclature *Murid herpesvirus 4*, MuHV-4), is a well-described small
70 animal model for the γHVs [4]. While these viruses establish a lifelong infection that is
71 often clinically inapparent, immune-suppressed individuals are particularly at risk for
72 γHV-associated malignancies [5].
73 Herpesvirus gene expression is extremely well-characterized in bulk populations.
74 Despite increasing evidence for single-cell heterogeneity in gene expression [6-8], there
75 remains limited understanding of herpesvirus infection at the single-cell level [9-12].
76 Here, we tracked endogenous viral and host RNAs using a sensitive, robust fluorescent
77 in situ hybridization assay combined with multiparameter flow cytometry (PrimeFlow™)
78 [13] to analyze the expression and inter-relationships of viral ncRNA, viral mRNA and
79 cellular mRNA at the single-cell level during γHV latency, reactivation and lytic
80 replication. These studies revealed unanticipated heterogeneity of infection,
81 emphasizing how single-cell analysis of virus infection can afford significant new
82 insights into the complexity of γHV infection.
83

84 **RESULTS**

85 *Single-cell analysis of viral RNAs during lytic infection.*

86 Traditional measurements of gene expression frequently rely on pooled cellular
87 material, obscuring intercellular variation in gene expression. To better define
88 expression of γHV RNAs at the single cell level, we employed the PrimeFlow™ RNA
89 assay [13] to study viral gene expression during murine gammaherpesvirus 68 (γHV68)
90 infection, a small animal γHV [4, 13]. This method is a highly sensitive, extremely
91 specific in situ hybridization assay, integrating Affymetrix-designed branched DNA
92 technology with single-cell analysis powered by multiparameter flow cytometry. This
93 method has been successfully used to detect both virus and host RNAs (e.g. in the
94 context of HIV infected individuals [13, 14]).

95 We first tested the ability of PrimeFlow™ to measure multiple viral RNAs during
96 lytic infection with γHV68, including small non-coding RNAs (tRNA-miRNA encoding
97 RNAs or TMERs [15]) and mRNAs. Mouse fibroblasts were infected with an
98 intermediate multiplicity of infection (MOI=5 plaque forming units of virus/cell) resulting
99 in a mixture of infected and uninfected cells. Under these conditions, TMER-5, one of
100 the eight γHV68 TMERs, and the γHV68 gene 73, were readily detectable by
101 conventional real-time PCR in γHV68-infected, but not mock-infected, cultures (Fig. 1A,
102 C). Parallel cultures were analyzed for RNA expression by PrimeFlow™. Whereas
103 mock-infected cells had no detectable expression of either the γHV68 TMERs or gene
104 73, WT γHV68-infected fibroblasts had a prominent population of TMER+ and gene 73+
105 cells, respectively (Fig. 1B and 1D). Infection of cells with a TMER-deficient γHV68
106 (TMER-TKO [16]), in which TMER expression is ablated through promoter disruption,

107 revealed no detectable TMER expression (Fig. 1B), yet robust gene 73 expression (Fig.
108 1D). Parallel studies revealed ready detection of gene 18, another γHV68 gene product
109 (Fig. 1E). These studies show that PrimeFlow™ is a sensitive, robust and specific
110 method to detect both viral non-coding and messenger RNAs during lytic infection,
111 quantifying both the frequency of gene expression and expression on a per cell basis.
112
113 *Heterogeneous gene expression during γHV latency and reactivation from latency.*
114 γHV latency is characterized by limited gene expression. We next
115 measured viral RNAs during latency and reactivation using the γHV68-infected
116 A20 HE2.1 cell line (A20.γHV68), a drug-selected latency model with restricted viral
117 gene expression that can reactivate following stimulation [17]. A20.γHV68 cells are
118 characterized by restricted viral gene expression, yet remain competent for reactivation
119 from latency and the production of infectious virions following chemical stimulation with
120 the phorbol ester, TPA [17, 18].
121 When we compared TMER expression between uninfected (parental, virus-
122 negative A20) and infected (A20.γHV68) cells by qRT-PCR, the viral ncRNA TMER-5
123 was exclusively detectable in A20.γHV68 cells, with minimal changes between
124 untreated and chemically-stimulated conditions (Fig. 2A). PrimeFlow™ analysis of
125 TMER expression in untreated A20.γHV68 cells revealed that a majority of these cells
126 expressed the TMERs, as defined by a positive signal in samples subjected to the
127 TMER probe relative to unstained cells (Fig. 2B). Untreated A20.γHV68 cells contained
128 a high frequency of cells expressing intermediate levels of TMERs (i.e. $TMER^{mid}$ cells),
129 with a significant signal enrichment above parental, virus-negative A20 cells (Fig. 2C).

130 While the frequency of TMER^{mid} cells remained relatively constant following TPA
131 stimulation (compare “Untreated” versus “Stimulated”, Fig. 2C), stimulated A20.γHV68
132 cultures also contained a small fraction of cells with high levels of TMERs (i.e. TMER^{high}
133 cells), not present in untreated cultures (Fig. 2C-D). Chemical stimulation is known to
134 result in variable penetrance of reactivation in latently infected cell lines [17]. Based on
135 this, we hypothesized that these rare, TMER^{high} cells may represent a subset of cells
136 that are undergoing reactivation from latency.

137 To test this, we analyzed the properties of TMER^{mid} and TMER^{high} cells,
138 comparing viral protein expression in untreated and stimulated A20.γHV68 cells. We
139 analyzed: 1) a γHV68 expressed GFP-hygromycin resistance fusion protein
140 (HygroGFP), under the control of a heterologous viral promoter (the CMV immediate
141 early promoter) [17], and 2) the γHV68 regulator of complement activation (RCA), a viral
142 protein encoded by the γHV68 gene 4 [19]. The vast majority of TMER^{mid} cells were
143 negative for HygroGFP and RCA (i.e. HygroGFP- RCA-), regardless of whether the
144 cells were present in untreated or stimulated cultures (Fig. 2E-F). Conversely, TMER^{high}
145 cells, which were uniquely present in stimulated cultures, had a significantly increased
146 frequency of HygroGFP+ cells with induction of RCA protein+ cells in a subset of cells
147 when compared to TMER^{mid} cells present in either untreated or stimulated cultures (Fig.
148 2E-F). By using imaging flow cytometry, we further analyzed the subcellular localization
149 of TMERs in TMER^{mid} cells compared with TMER^{high} RCA+ cells. TMERs were
150 predominantly nuclear in both TMER^{mid} and TMER^{high} RCA+ cells, as defined by co-
151 localization with DAPI fluorescence (Fig. 2G). These data demonstrate that the TMERs
152 are expressed during latency, and that following reactivation-inducing stimulation,

153 TMERs are further induced in a rare subset of cells which are characterized by
154 increased viral transcription and translation.
155
156 *Detection of endogenous viral gene expression during KSHV latency and reactivation.*
157 To extend these findings, we analyzed viral gene expression in the KSHV
158 infected B cell tumor line, BCBL-1, focused on detection of an abundant viral ncRNA,
159 the KSHV polyadenylated nuclear RNA (PAN, nut1, or T1.1) [20]. PAN RNA is known to
160 be highly inducible upon induction of reactivation in KSHV latently infected B cell
161 lymphoma cell lines [10, 20]. The frequency of PAN RNA+ cells was low in untreated
162 BCBL-1 cells, with ~1% of cells spontaneously expressing this ncNRA (Fig. 3A-B).
163 Despite the low frequency, this hybridization was clearly above background, as ncRNA
164 defined on the KSHV- and EBV-negative B cell lymphoma cell line BL41 [21, 22] (Fig.
165 3A-B). Upon stimulation of BCBL-1 cells with the reactivation-inducing stimuli TPA and
166 sodium butyrate, the frequency of PAN RNA+ cells significantly increased with
167 expression in ~25% of cells (Fig. 3A-B). Although stimulation of BCBL-1 cells
168 significantly increased the frequency of PAN RNA+ events compared to untreated
169 cultures, PAN RNA expression on an individual cell basis was comparable between
170 cells from untreated or stimulated cultures (Fig. 3C). As anticipated, stimulation of
171 BCBL-1 cells was associated with increased viral DNA, consistent with stimulated
172 cultures undergoing reactivation from latency (Supplemental Fig. 1).
173 We next analyzed the properties of BCBL-1 cells as a function of PAN RNA
174 expression. In untreated cells, PAN RNA+ or RNA- cells had comparable cell size
175 (define by forward scatter, FSC) and granularity (defined by side scatter, SSC). Gene

176 73 expression was low in untreated BCBL-1 samples, with signal intensity in PAN RNA-
177 cells close to the background fluorescence observed in unstained samples. PAN RNA+
178 cells in untreated cultures had a modest increase in gene 73 expression relative to PAN
179 RNA- cells (Fig. 3D-E). In stimulated BCBL-1 cultures, PAN RNA+ cells had a modest
180 decrease in cell size (defined by forward scatter) and a trend towards reduced
181 granularity (defined by side scatter) compared to PAN RNA- cells (Fig. 3F-G).
182 Stimulated BCBL-1 cultures also had an increased gene 73 signal when compared to
183 unstained samples (Fig. 3F), with PAN RNA+ cells again showing ~2-fold increase
184 compared to PAN RNA- cells (Fig. 3F-G). These data demonstrate robust detection of
185 PAN RNA by PrimeFlow™, and further identify PAN RNA expression in a subset of both
186 untreated and reactivation-induced BCBL-1 cells.
187
188 *Detection of endogenous viral gene expression during EBV latency and reactivation.*
189 EBV encodes two abundant non-coding RNAs, the EBV-encoded RNAs (EBERs)
190 EBER1 and EBER2. We tested the ability of the PrimeFlow™ method to detect EBER in
191 an EBV positive, Burkitt lymphoma type I latency cell line, Mutu I [23]. EBER expression
192 was detected in a ~45% of Mutu I cells in either untreated or stimulated conditions, with
193 no significant probe hybridization in the KSHV- and EBV-negative BL41 cell line (Fig.
194 4A). Stimulated Mutu I cells showed a modest, 2-fold increase in EBER expression on
195 an individual cell basis, relative to untreated EBER+ cells (Fig. 4B). Based on these
196 data, EBER expression in Mutu I cells appears to be constitutive, with stimulation under
197 these conditions resulting in minimal consequences on either the frequency or per-cell
198 expression of the EBERs.

199

200 *Single-cell analysis of actin mRNA degradation as a readout of virus-induced host*

201 *shutoff.*

202 Many herpesviruses, including γHV68, EBV and KSHV, induce host shutoff

203 during lytic replication and reactivation from latency, a process characterized by

204 dramatic decreases in host mRNAs [24, 25]. Consistent with published reports [24],

205 qRT-PCR analysis of a cellular housekeeping gene, β-actin (Actb), showed reduced

206 actin mRNA in γHV68 lytically-infected fibroblasts by 18 hours pi (Fig. 5A). While mock-

207 infected samples had a uniformly positive population of actin RNA^{high} cells detectable by

208 PrimeFlow^{TM}, γHV68-infected fibroblast cultures demonstrated a bimodal distribution of

209 actin RNA^{high} and actin RNA^{low} cells (Fig. 5B). The actin RNA^{low} population had a

210 fluorescent signal that was only modestly above background fluorescence (defined by

211 the "No probe" sample), suggesting an all-or-none phenomenon in which cells either

212 had no change in actin RNA levels or had pronounced actin RNA degradation.

213 Simultaneous analysis of TMER and actin RNA expression revealed that actin RNA^{low}

214 cells were frequently TMER^{high}, with actin^{high} cells frequently TMER^{negative} at this time

215 (Fig. 5C).

216 To determine whether actin RNA regulation could also be observed during γHV68

217 latency and reactivation, we measured actin RNA levels in A20.γHV68 cells. Parental,

218 virus-negative A20 cells and A20.γHV68 cells had relatively comparable actin RNA

219 levels by qRT-PCR, in both untreated and stimulated cells (Fig. 5D). Given that host

220 shutoff is expected to primarily occur in rare, reactivating cells, we measured actin RNA

221 degradation relative to TMER expression by the PrimeFlow^{TM} method. Untreated

222 A20.γHV68 cultures had no discernable population of TMER+ actin RNA^{low} events,
223 whereas stimulated cultures were characterized by a rare population of TMER^{high} actin
224 RNA^{low} cells (Fig. 5E). We further compared actin RNA expression between TMER^{mid}
225 and TMER^{high} cells, in untreated versus stimulated cultures using our previously defined
226 subpopulations (Fig. 2). While TMER^{mid} cells from either untreated or stimulated
227 cultures were predominantly actin RNA+, TMER^{high} cells from stimulated cultures
228 showed a significant increased frequency of actin RNA^{low} events (Fig. 5F-G). These
229 studies reveal actin RNA as a sensitive indicator of virus-induced host shutoff, and
230 demonstrate this as an all-or-none phenomenon that can be readily queried at the
231 single-cell level.
232
233 *Heterogeneity of gene expression during de novo lytic replication.*
234 Next, we revisited our analysis of gene expression during de novo lytic infection
235 of fibroblasts, to examine co-expression relationships between viral ncRNA (TMERs),
236 viral mRNA (the γHV68 gene 73), viral protein (RCA protein) and cellular actin mRNA
237 degradation [19, 24]. Fibroblast cultures were infected with an intermediate multiplicity
238 of infection to produce a mixture of uninfected and infected cells, and then subjected to
239 the PrimeFlow™ method.
240 To enable an unbiased, automated analysis of gene expression profiles in γHV68
241 lytically infected cells relative to mock infected cells, data were subjected to the
242 automated clustering algorithm X-shift [26], to identify potential subpopulations of cells
243 with heterogeneous gene expression in these cultures. By sampling 1,000 cells from
244 multiple mock- and virus-infected cultures, the X-shift algorithm consistently identified 7

245 major clusters of cells (Fig. 6A) defined by varying gene expression patterns. While
246 some of the clusters were exclusively found in mock-infected cultures, virus-infected
247 cultures contained three broad types of cell clusters: 1) uninfected cells, with no viral
248 gene expression and normal actin RNA, 2) fully infected cells, with robust expression of
249 the TMERs, gene 73, actin RNA downregulation and frequent expression of the RCA
250 protein, and 3) intermediate populations characterized by variable expression of the
251 TMERs and gene 73 (Fig. 6A).

252 To validate these findings using a more conventional method, we compared
253 TMER and gene 73 RNA co-expression on a biaxial plot. By comparing mock-infected,
254 WT-infected and TMER-TKO-infected cultures, this analysis revealed five populations of
255 gene expression (Fig. 6B), including cells with: 1) no detectable expression of either
256 viral RNA (TMER- gene 73-), 2) TMER+ gene 73- cells (bottom right quadrant), 3)
257 TMER- gene 73+ cells (upper left quadrant), 4) $TMER^{low}$ gene 73^{low} cells (lower left
258 edge of the upper right quadrant), and 5) $TMER^{high}$ gene 73^{high} cells (upper right
259 quadrant). The definition of TMER positive events was defined based on background
260 fluorescent levels observed in TMER-TKO infected cultures (Fig. 6B). These 5
261 populations were each assigned a unique color for subsequent analysis (Fig. 6C).

262 We then compared the cellular phenotype and gene expression within these 5
263 distinct populations. Analysis of TMERs, gene 73, actin RNA, RCA protein, cell size
264 (forward scatter), and granularity (side scatter) revealed multiple types of viral gene
265 expression. As expected, TMER- gene 73- cells (in black) had no evidence of virus
266 infection, with no detectable viral protein (RCA) or actin downregulation (Fig. 6D-E).
267 Cells with low expression of either the TMERs and/or gene 73 contained viral RNAs, but

268 had minimal expression of either viral protein or actin downregulation (Fig. 6D-E). In
269 stark contrast, cells that were TMER^{high} gene 73^{high} (in red, Fig. 6D-E) had multiple
270 characteristics of progressive virus infection including a prominent fraction of cells that
271 expressed RCA and/or had actin RNA downregulation. Further, TMER^{high} gene 73^{high}
272 cells were consistently smaller in cell size (defined by forward scatter, FSC) and higher
273 in granularity (defined by side scatter, SSC), a feature that was unique to this phenotype
274 (Fig. 6D-E).

275 Given the heterogeneous patterns of RNA and protein expression among
276 lytically-infected cells, we next queried TMER subcellular localization as a function of
277 viral gene expression using imaging flow cytometry. While the majority of TMER+ cells
278 had a primarily nuclear TMER localization (defined by DAPI co-localization, as in [27]),
279 the frequency of cells with nuclear TMER localization was highest among TMER+ gene
280 73- cells and lowest among TMER+ gene 73+ RCA+ cells (Fig. 7A-B). These data
281 suggest that the TMERs can be localized in either the nucleus or cytoplasm during
282 γHV68 lytic replication, and that this localization is not strictly a function of magnitude of
283 gene expression.

284 Finally, we used tSNE, a dimensionality reduction algorithm, to better delineate
285 the relationship between TMER, gene 73, RCA protein and actin downregulation across
286 populations defined by variable TMER and gene 73 expression. Consistent with our
287 histogram analysis (Fig. 6D-E), uninfected and intermediate populations that expressed
288 either TMERs or gene 73 were relatively uniform in gene expression (Fig. 7C). In
289 contrast, TMER^{high} gene 73^{high} cells expressed a wider array of phenotypes, including
290 both a predominant fraction of cells that were actin RNA^{low} RCA+, as well as a distinct

291 group of cells that were actin RNA+ RCA- (Fig. 7C). Notably, RCA expression and actin
292 degradation were inversely correlated, with very few cells that expressed RCA also high
293 for actin RNA. Actin RNA+ populations among TMER^{high} gene 73^{high} cells were
294 associated with larger cell size (Supplementary Fig. 2). The diversity of phenotypes
295 among TMER^{high} gene 73^{high} cells was confirmed by biaxial gating of actin RNA versus
296 RCA protein expression (Fig. 7D). In total, these data indicate heterogeneous
297 progression of lytic replication in vitro. While some cells have robust viral mRNA and
298 protein expression, additional cell subsets are characterized by limited or divergent
299 gene expression.

300 **DISCUSSION**

301 Herpesvirus gene expression has been historically analyzed in bulk cell
302 populations. These studies have provided an essential cornerstone to understanding
303 the transcriptional and translational capacity of the herpesviruses. Despite this, recent
304 studies on cellular and viral transcription from other systems have emphasized a high
305 degree of cell-to-cell variation in gene expression [6-11, 13, 14], something we have
306 further investigated here. By applying the PrimeFlow^{TM} methodology to measure
307 endogenous viral gene expression across multiple gammaherpesviruses, and multiple
308 stages of infection, we have gained critical new insights into the inter-relationships of
309 gene expression at the single-cell level.

310 A primary focus of the current study has been to analyze expression of γHV
311 ncRNAs. Although the TMERs, EBERs and PAN RNA all represent abundant γHV
312 ncRNAs, these ncRNAs are transcribed by distinct mechanisms: KSHV PAN is a highly-
313 inducible, RNA pol II-transcribed ncRNA [20], in contrast to the RNA pol III-transcribed
314 TMERs and EBERs [15, 28]. This differential regulation was mirrored in the expression
315 patterns we observed. Whereas TMERs and EBERs were detected in a large fraction of
316 latently infected cells, PAN RNA was expressed in a low frequency of latently infected
317 cells, with prominent induction following cell stimulation and the induction of
318 reactivation. The viral ncRNAs were efficiently detected, as might be predicted due to
319 their abundance. The viral gene 73 encodes a transcription factor that is expressed at a
320 far lower level and are also efficiently detected, demonstrating that rare mRNAs can be
321 measured coincidently with abundant RNAs and with proteins, with no modifications
322 required. A unique advantage of our current approach is the ability to measure the

323 frequency of ncRNA expressing cells and changes in expression within individual cells.
324 This has been particularly insightful for the identification of rare PAN RNA+ cells in
325 untreated BCBL-1 cells and a TMER^{high} subpopulation of cells in reactivating
326 A20.γHV68 cells. Integrating this method with cell sorting will afford future opportunities
327 to investigate unique properties of these rare cell populations.

328 Among the viral ncRNAs measured, in-depth analysis of TMER expression
329 during γHV68 infection has revealed new insights into infection. In the context of
330 latency, the TMERs are constitutively expressed in many, but not all, latently infected
331 cells using the A20.γHV68 model. Further, stimulating these cells to undergo
332 reactivation has a minimal effect on the frequency of cells expressing intermediate
333 levels of TMERs (i.e. TMER^{mid} cells), instead resulting in the appearance of a minor
334 population of TMER^{high} cells. Notably, TMER^{high} cells show additional features of lytic
335 cycle progression, including actin RNA degradation and RCA protein expression. Why
336 only some latently infected cells show the TMER^{high} phenotype, and what regulates the
337 inducible expression of the RNA pol III-transcribed TMERs remain important questions
338 raised by this analysis.

339 Of the γHVs studied here, only γHV68 has a robust in vitro lytic replication
340 system. Our studies on γHV68 lytic replication revealed multiple unanticipated results.
341 By using cultures that contained both infected and uninfected cells, our analysis
342 identified at least four different subsets of cells, stratified by differential viral gene
343 expression of the TMERs and gene 73. Strikingly, during lytic replication there were
344 many cells with limited viral gene expression, expressing low levels of either the TMERs
345 and/or gene 73, but lacking additional signs of virus gene expression (i.e. actin RNA

346 degradation or RCA protein expression). Conversely, only some viral RNA+ cells
347 showed robust viral expression characterized by a constellation of gene expression,
348 defined as TMER^{high} gene 73^{high} actin RNA^{low} RCA+. While there is precedence that
349 reactivation from latency in KSHV infection can be asynchronous [9], this heterogeneity
350 of viral gene expression during in vitro lytic replication was unanticipated and suggests
351 that lytic infection under these reductionist conditions is either asynchronous, abortive,
352 or inefficient. This heterogeneity of gene expression raises important questions
353 regarding the universality of the prototypical cascade of immediate early, early and late
354 gene expression that is widely accepted in the herpesvirus field and suggests additional
355 levels of complexity that may be obscured by bulk cell analysis.
356 This method allows multiplexed analysis of single-cell gene expression, to both
357 directly measure viral RNAs and downstream consequences of gene expression
358 including viral protein production and host RNA degradation, secondary to protein
359 translation. This approach has notable advantages to conventional analyses of gene
360 expression: 1) it can measure endogenous viral gene expression (both mRNA and
361 ncRNA) in the absence of recombinant viruses or marker genes, and 2) it can rapidly
362 analyze gene and protein expression inter-relationships, across millions of cells,
363 providing unique complementary strengths to other single-cell methodologies (e.g.
364 single-cell RNA-seq). In future, this method can be further integrated with additional
365 antibody-based reagents, to simultaneously query post-translational modifications (e.g.
366 protein phosphorylation) as a function of cell cycle stage. It is also notable that through
367 the use of imaging flow cytometry, it is possible to interrogate subcellular RNA and
368 protein localization throughout distinct stages of virus infection.

369 In total, these studies demonstrate the power of single-cell analysis of
370 herpesvirus gene expression. Our data emphasize the heterogeneity of γHV gene
371 expression at the single-cell level. The factors that underlie this heterogeneity are
372 currently unknown, but could reflect either asynchronous or inefficient infection in many
373 infected cells (e.g. in the context of lytic infection). Whether this variation arises from
374 viral or cellular heterogeneity is a fundamental question for future research.

375 **MATERIALS AND METHODS**

376 Viruses and tissue culture. All γHV68 viruses were derived from the γHV68 strain
377 WUMS (ATCC VR-1465) [29], using bacterial artificial chromosome-derived wild-type
378 (WT) γHV68 or γHV68.TMER-Total KnockOut (TMER-TKO) [16]. Virus stocks were
379 passaged, grown, and titered as previously described [16]. Mouse 3T12 fibroblasts
380 (ATCC CCL-164) were inoculated with 5 plaque forming units/cell for one hour, followed
381 by inoculum removal and replacement with fresh media, with analysis between 8-18 hpi.
382 Parental, virus-negative A20 B cells, or A20.γHV68 (HE2.1) B cells [17], were treated
383 with vehicle (untreated) or stimulated with 12-O-tetradecanolphrobol-13-actate (TPA) 20
384 ng/ml (Sigma) (in DMSO) harvested 24 hr later. BCBL-1 B cells, latently infected with
385 KSHV (HHV-8), were cultured in RPMI containing 20% FBS, 1% Penicillin/Streptomycin
386 with L-glutamine, 1% HEPES and 50 μM βME. BCBL-1 B cells were treated with vehicle
387 (untreated) or stimulated with 20 ng/ml TPA (in DMSO) and Sodium Butyrate (NaB) 0.3
388 mM (Calbiochem) (in water) and then harvested 72 hr later. BL41 B cells (negative for
389 KSHV and EBV) were cultured in RPMI with 10% FBS, 1% Penicillin/Streptomycin with
390 L-glutamine, and 50 μmol βME. Mutu I cells, an EBV-infected, type I latency Burkitt's
391 lymphoma cell line [23] were cultured in RPMI with 10% FBS, 1%
392 Penicillin/Streptomycin and L-glutamine. Mutu I or BL41 B cells were either treated with
393 vehicle (DMSO) or stimulated with 20 ng/ml TPA (in DMSO) and then harvested 48 hr
394 later. BCBL-1 and BL41 cells were generously provided by Dr. Rosemary Rochford
395 (University of Colorado), with additional BCBL-1 cells obtained from the NIH AIDS
396 reagent program (catalog # 3233). Mutu I cells were generously provided by Dr.
397 Shannon Kenney (University of Wisconsin).

398

399 Flow cytometric analysis. Cells were harvested at the indicated time points and
400 processed for flow cytometry using the PrimeFlow™ RNA Assay (Thermo Fisher). Cells
401 were incubated with an Fc receptor blocking antibody (2.4G2) for 10 min and then fixed
402 with 2% PFA (Fisher), washed with PBS (Life Technology). Cells were stained with a
403 rabbit antibody against the γHV68 gene 4 protein, regulator of complement activation
404 (RCA) [19], labeled with Zenon R-phycoerythrin rabbit IgG label reagent (Life
405 Technologies) following manufacturer's protocol. Samples were subjected to the
406 PrimeFlow™ RNA Assay following manufacturer's protocols, using viral and host target
407 probes conjugated to fluorescent molecules (Table S1). DAPI (BioLegend) was used on
408 a subset of samples following manufacturer's protocol, prior to PrimeFlow™ probe
409 hybridization. Flow cytometric analysis was done on LSR II (BD Biosciences), Fortessa
410 (BD Biosciences), and ZE5 (Bio-Rad) flow cytometers, with compensation values based
411 on antibody-stained beads (BD Biosciences), and modified as needed post-collection
412 using FlowJo.

413

414 Imaging flow cytometry. Cells were treated as described above then harvested, and split
415 into two aliquots: one for conventional flow cytometry, and one for imaging flow
416 cytometry, acquired on an Amnis ImageStream®X Mark II imaging flow cytometer
417 (MilliporeSigma) with a 60X objective and low flow rate/high sensitivity using INSPIRE®
418 software. Brightfield (BF) and side scatter (SSC) images were illuminated by LED light
419 and a 785nm laser respectively. Fluorescent probes were excited off 405nm, 488nm,
420 and 642nm lasers with the power adjusted properly to avoid intensity saturation of the

421 camera. Single color controls for compensation were acquired by keeping the same
422 acquisition setting for samples, with the difference of turning the BF LED light and
423 785nm (SSC) laser off.
424 The acquired data were analyzed using IDEAS® software (MilliporeSigma).
425 Single cells that were in focus were defined as a population with a high "gradient RMS"
426 value, an intermediate "Area" value, and a medium to high "Aspect ratio" value for
427 subsequent analysis. Positive and negative events for each fluorescent marker were
428 determined using the "Intensity" feature. TMER nuclear localization was quantified using
429 "Similarity" feature, the log-transformed Pearson's correlation coefficient by analyzing
430 the pixel values of two image pairs [27]. The degree of nuclear localization of TMER
431 was measured by correlating the pixel intensity of two images with the same spatial
432 registry. The paired TMER and DAPI images were quantified by measuring the
433 "Similarity Score" which cells with high similarity scores display high TMER nuclear
434 localization with similar image pairs. By contrast cells with low similarity scores show
435 low TMER nuclear localization with dissimilar image pairs.
436
437 RNA and qRT-PCR. RNA was isolated using Trizol (Life Technologies) per
438 manufacturer's protocol and re-suspended in DEPC treated water. 3 μg of RNA was
439 treated with DNase 1 (Promega) for 2 hours at 37°C, heat inactivated for 10 min at
440 65°C. 500 ng of RNA was then subjected to reverse transcription using SuperScript II
441 (Life Technologies) following manufacturer's protocol for gene specific, oligo(dT), or
442 random primers (Life Technologies). Quantitative PCR (qPCR) was performed using iQ
443 SYBR Green super mix (Bio-Rad) follow manufacturer's protocol using host and viral

444 primer sets (Table S2) or using QuantiTech Primer Assay (Qiagen) for 18s (Hs-
445 RRN18S_1_SG). qPCR conditions: 3 min at 95°C, amplification cycles for 40 cycles of
446 15 sec at 95°C, annealing/ extension at temperature for specific primer set for 1 min
447 ending with a melt curve which started at 50°C or 55°C to 95°C increasing 0.5°C for
448 0:05 sec. A standard curve for each primer set was generated by pooling a portion of
449 each sample together and doing a 1:3 serial dilution. 75 ng of cDNA of the unknown
450 samples was loaded per qPCR reaction/primer set, with reactions run on a Bio-Rad 384
451 CFX LightCycler and data analyzed using Bio-Rad CFX manager software. Data
452 analysis was done using the 1:3 standard curve as the control Ct value to calculate the
453 delta ct, and the Pfaffl equation was used to define the fold difference between the gene
454 of interest and 18s (reference gene) [30]. qPCR products were analyzed by melt curve
455 analysis, with all reactions having a prominent, uniform product. In the case of primers
456 with an aberrant melt curve product (e.g. that arose at late cycles), products were
457 clearly a different product as defined by melt curve analysis.
458
459 KSHV genome quantification. BCBL-1 or BL41 cells were plated at 7.5e5 cells/well in a
460 6 well plate with 20 ng/ml TPA and 0.3 mM NaB or vehicle only (DMSO and H_2O). Cells
461 and supernatant were harvested at 72 hrs post-treatment, hard-spun for 30 min at 4° C
462 and DNA was isolated using the DNeasy Blood and Tissue kit following manufacturer's
463 protocol, except for sample digestion for 1 hour instead of 10 min. 100 ng of DNA per
464 sample was used for qPCR analysis via SYBR green detection using KSHV ORF50
465 primers (5' -TCC GGC GGA TAT ACC GTC AC- 3' and 5'- GGT GCA GCT GGT ACA
466 GTG TG-3') [31]. qPCR was analyzed using relative quantification normalized against

467 unit mass calculation, ratio = $E^{deltaCt}$ (Real-Time PCR Application Guide, Bio-Rad
468 Laboratories Inc. 2006).
469
470 Software and Statistical analysis. All flow cytometry data were analyzed in FlowJo
471 (version 8.8.7 or 10.5.0), with flow cytometry data shown either as histogram overlays or
472 pseudo-color dot plots (with or without smoothing), showing outliers (low or high
473 resolution) on $\log_{10}$ scales. Statistical analysis and graphing were done in GraphPad
474 Prism (Version 6.0d and 7.0d). Statistical significance was tested by unpaired t test
475 (when comparing two conditions) or by one-way ANOVA (when comparing three or
476 more samples) subjected to multiple corrections tests using recommended settings in
477 Prism. *X-shift analysis:* For automated mapping of flow cytometry data using X-shift,
478 data were obtained from compensated flow cytometry files, exported from FlowJo, using
479 singlets that were live (defined by sequential gating on single cells by FSC-H vs. FSC-W
480 and SSC-H vs. SSC- W, that were DAPI bright vs. SSC-A). These events were imported
481 into the Java based program VorteX (http://web.stanford.edu/~samusik/vortex/) [26].
482 Four parameters [TMER (AlexaFluor (AF) 488), RCA (PE), gene 73 (AF647), and Actin
483 (AF750)] were selected for clustering analysis using the X-shift algorithm. The following
484 settings were used when importing the data set into VorteX: i) Numerical transformation:
485 arcsinh(x/f), f=150, ii) noise threshold: apply noise threshold of 1.0 (automatic and
486 recommended setting), iii) feature rescaling: none, and iv) normalization: none, v) a
487 Euclidean noise filter was used with a Minimal Euclidean length of the profile of 1.0, and
488 vi) an import max of 1,000 rows from each file after filtering was selected. The following
489 settings were used when preparing the data set for clustering analysis: i) distance

490 measure: angular distance, ii) clustering algorithm: X-shift (gradient assignment), iii)
491 density estimate: N nearest neighbors (fast), iv) number of neighbors for density
492 estimate (K): from 150 to 5, with 30 steps, and v) number of neighbors for mode finding
493 (N): determine automatically. After the cluster analysis was completed, all results were
494 selected and the K value that corresponded with optimal clustering (the elbow point)
495 was calculated, in this case K= 50. All clusters (seven clusters total) for the optimal K
496 value were selected and a Force-Directed Layout was created. The maximum number
497 of events sampled from each cluster was 20, and the number of nearest neighbors was
498 10. All settings used for this analysis were automated or explicitly recommended
499 (https://github.com/nolanlab/vortex/wiki). Force-Directed layouts in Fig. 6A were saved
500 as graphml files from VorteX, opened in the application Gephi v 0.9.1, and colored by
501 different variables (Cluster ID, experimental group, Actin mRNA, RCA, Gene 73, and
502 TMERs respectively) in Adobe Illustrator CC 2017. Full details on use of the X-shfit
503 algorithm and analysis pipeline can be found in [32]. *tSNE analysis:* Gated events for
504 each of the six identified populations were exported from FlowJo, and then imported into
505 Cytobank (www.cytobank.org) for analysis using the viSNE algorithm. Each file was
506 used for a separate viSNE analysis (six total runs), where all available events were
507 selected for clustering (202,669, 128,028, 29,096, 5,610, 8,956, 35,850 respectively)
508 and four parameters were selected for clustering (Actin mRNA, RCA, Gene 73, and
509 TMERs). The resulting tSNE plots were colored according to expression using the
510 "rainbow" color option, with individual events shown using the stacked dot option. The
511 channel range was user-defined for each marker according to the range in expression
512 established in Fig. 6E.
513

514 **ACKNOWLEDGMENTS**

515 The authors would like to acknowledge insightful comments made by members of the
516 Clambey and van Dyk laboratories, support of flow cytometry services through
517 ClinImmune, the Dept. of Immunology & Microbiology, and the University of Colorado
518 Cancer Center, the provision of KSHV and EBV cell lines by Dr. Rosemary Rochford
519 (University of Colorado) and Dr. Shannon Keeney (University of Wisconsin), and expert
520 technical guidance from Matt Cato and Dr. Nori Ueno (Thermo Fisher). T.C. and B.A.
521 are employees of EMD Millipore and have a potential conflict of interest. E.T.C. was a
522 recipient of the 2016 North America Affymetrix Single-cell Grant Recipient, for studies
523 unrelated to this manuscript.

524

525 **FUNDING**

526 This research was funded by National Institutes of Health grants R01CA103632 and
527 R01CA168558 to L.F.V.D., R21AI134084 to E.T.C. and L.F.V.D., and by an American
528 Heart Association National Scientist Development grant (#13SDG14510023), a
529 Colorado CTSI Novel methods development grant, and funding from the University of
530 Colorado Dept. of Anesthesiology to E.T.C.. The Colorado CTSI is supported by
531 NIH/NCATS Colorado CTSA Grant Number UL1 TR002535. Contents are the authors'
532 sole responsibility and do not necessarily represent official NIH views. The funders had
533 no role in study design, data collection and analysis, decision to publish, or preparation
534 of the manuscript.

535

AUTHOR CONTRIBUTIONS

Conceptualization: Eric T. Clambey, Linda F. van Dyk

Data curation: Lauren M. Oko, Eric T. Clambey

Formal analysis: Lauren M. Oko, Abigail K. Kimball, Rachael E. Kaspar, Benjamin Alderete, Tim Chang, Eric T. Clambey

Funding acquisition: Eric T. Clambey, Linda F. van Dyk

Investigation: Lauren M. Oko, Ashley N. Knox

Methodology: Lauren M. Oko, Benjamin Alderete, Tim Chang, Linda F. van Dyk, Eric T. Clambey

Project administration: Eric T. Clambey, Linda F. van Dyk

Resources: Lauren M. Oko, Benjamin Alderete, Tim Chang

Software: Abigail K. Kimball

Supervision: Eric T. Clambey, Linda F. van Dyk

Validation: Lauren M. Oko, Eric T. Clambey

Visualization: Lauren M. Oko, Abigail K. Kimball, Rachael E. Kaspar, Benjamin Alderete, Tim Chang, Eric T. Clambey

Writing – original draft: Eric T. Clambey, Linda F. van Dyk

Writing – review & editing:

555 **REFERENCES**

556

557 1. Pellett PJ, Roizman B. *Herpesviridae*. Fields Virology. 6th ed2013.

558 2. Longnecker RM, Kieff, E., Cohen, J.I. Epstein-Barr Virus. Fields Virology. 6th

559 ed2013.

560 3. Ganem D. KSHV infection and the pathogenesis of Kaposi's sarcoma. Annu Rev

561 Pathol. 2006;1:273-96. doi: 10.1146/annurev.pathol.1.110304.100133. PubMed PMID:

562 18039116.

563 4. Barton E, Mandal P, Speck SH. Pathogenesis and host control of

564 gammaherpesviruses: lessons from the mouse. Annual review of immunology.

565 2011;29:351-97. doi: 10.1146/annurev-immunol-072710-081639. PubMed PMID:

566 21219186.

567 5. Cesarman E. Gammaherpesviruses and lymphoproliferative disorders. Annu Rev

568 Pathol. 2014;9:349-72. doi: 10.1146/annurev-pathol-012513-104656. PubMed PMID:

569 24111911.

570 6. Shalek AK, Satija R, Adiconis X, Gertner RS, Gaublomme JT, Raychowdhury R,

571 et al. Single-cell transcriptomics reveals bimodality in expression and splicing in immune

572 cells. Nature. 2013;498(7453):236-40. doi: 10.1038/nature12172. PubMed PMID:

573 23685454; PubMed Central PMCID: PMCPMC3683364.

574 7. Battich N, Stoeger T, Pelkmans L. Control of Transcript Variability in Single

575 Mammalian Cells. Cell. 2015;163(7):1596-610. doi: 10.1016/j.cell.2015.11.018. PubMed

576 PMID: 26687353.

577 8. Chen KH, Boettiger AN, Moffitt JR, Wang S, Zhuang X. RNA imaging. Spatially

578 resolved, highly multiplexed RNA profiling in single cells. Science.

579 2015;348(6233):aaa6090. doi: 10.1126/science.aaa6090. PubMed PMID: 25858977;
580 PubMed Central PMCID: PMCPMC4662681.
581 9. Adang LA, Parsons CH, Kedes DH. Asynchronous progression through the lytic
582 cascade and variations in intracellular viral loads revealed by high-throughput single-cell
583 analysis of Kaposi's sarcoma-associated herpesvirus infection. J Virol.
584 2006;80(20):10073-82. doi: 10.1128/JVI.01156-06. PubMed PMID: 17005685; PubMed
585 Central PMCID: PMCPMC1617294.
586 10. Borah S, Nichols LA, Hassman LM, Kedes DH, Steitz JA. Tracking expression
587 and subcellular localization of RNA and protein species using high-throughput single
588 cell imaging flow cytometry. RNA. 2012;18(8):1573-9. doi: 10.1261/rna.033126.112.
589 PubMed PMID: 22745225; PubMed Central PMCID: PMCPMC3404377.
590 11. Ma JZ, Russell TA, Spelman T, Carbone FR, Tscharke DC. Lytic gene
591 expression is frequent in HSV-1 latent infection and correlates with the engagement of a
592 cell-intrinsic transcriptional response. PLoS Pathog. 2014;10(7):e1004237. doi:
593 10.1371/journal.ppat.1004237. PubMed PMID: 25058429; PubMed Central PMCID:
594 PMCPMC4110040.
595 12. Shnayder M, Nachshon A, Krishna B, Poole E, Boshkov A, Binyamin A, et al.
596 Defining the Transcriptional Landscape during Cytomegalovirus Latency with Single-
597 Cell RNA Sequencing. MBio. 2018;9(2). Epub 2018/03/15. doi: 10.1128/mBio.00013-18.
598 PubMed PMID: 29535194; PubMed Central PMCID: PMCPMC5850328.
599 13. Porichis F, Hart MG, Griesbeck M, Everett HL, Hassan M, Baxter AE, et al. High-
600 throughput detection of miRNAs and gene-specific mRNA at the single-cell level by flow

601 cytometry. Nat Commun. 2014;5:5641. doi: 10.1038/ncomms6641. PubMed PMID:
602 25472703; PubMed Central PMCID: PMCPMC4256720.
603 14. Baxter AE, Niessl J, Fromentin R, Richard J, Porichis F, Charlebois R, et al.
604 Single-Cell Characterization of Viral Translation-Competent Reservoirs in HIV-Infected
605 Individuals. Cell Host Microbe. 2016;20(3):368-80. doi: 10.1016/j.chom.2016.07.015.
606 PubMed PMID: 27545045; PubMed Central PMCID: PMCPMC5025389.
607 15. Diebel KW, Claypool DJ, van Dyk LF. A conserved RNA polymerase III promoter
608 required for gammaherpesvirus TMER transcription and microRNA processing. Gene.
609 2014;544(1):8-18. doi: 10.1016/j.gene.2014.04.026. PubMed PMID: 24747015.
610 16. Diebel KW, Oko LM, Medina EM, Niemeyer BF, Warren CJ, Claypool DJ, et al.
611 Gammaherpesvirus small noncoding RNAs are bifunctional elements that regulate
612 infection and contribute to virulence in vivo. MBio. 2015;6(1):e01670-14. doi:
613 10.1128/mBio.01670-14. PubMed PMID: 25691585; PubMed Central PMCID:
614 PMCPMC4337559.
615 17. Forrest JC, Speck SH. Establishment of B-cell lines latently infected with
616 reactivation-competent murine gammaherpesvirus 68 provides evidence for viral
617 alteration of a DNA damage-signaling cascade. J Virol. 2008;82(15):7688-99. doi:
618 10.1128/JVI.02689-07. PubMed PMID: 18495760; PubMed Central PMCID:
619 PMCPMC2493333.
620 18. Cheng BY, Zhi J, Santana A, Khan S, Salinas E, Forrest JC, et al. Tiled
621 microarray identification of novel viral transcript structures and distinct transcriptional
622 profiles during two modes of productive murine gammaherpesvirus 68 infection. J Virol.

623 2012;86(8):4340-57. doi: 10.1128/JVI.05892-11. PubMed PMID: 22318145; PubMed
624 Central PMCID: PMCPMC3318610.
625 19. Kapadia SB, Molina H, van Berkel V, Speck SH, Virgin HWt. Murine
626 gammaherpesvirus 68 encodes a functional regulator of complement activation. J Virol.
627 1999;73(9):7658-70. PubMed PMID: 10438856; PubMed Central PMCID: PMC104293.
628 20. Sun R, Lin SF, Gradoville L, Miller G. Polyadenylylated nuclear RNA encoded by
629 Kaposi sarcoma-associated herpesvirus. Proceedings of the National Academy of
630 Sciences of the United States of America. 1996;93(21):11883-8. Epub 1996/10/15.
631 PubMed PMID: 8876232; PubMed Central PMCID: PMCPMC38153.
632 21. Marchini A, Longnecker R, Kieff E. Epstein-Barr virus (EBV)-negative B-
633 lymphoma cell lines for clonal isolation and replication of EBV recombinants. J Virol.
634 1992;66(8):4972-81. Epub 1992/08/01. PubMed PMID: 1321281; PubMed Central
635 PMCID: PMCPMC241347.
636 22. Cao S, Strong MJ, Wang X, Moss WN, Concha M, Lin Z, et al. High-throughput
637 RNA sequencing-based virome analysis of 50 lymphoma cell lines from the Cancer Cell
638 Line Encyclopedia project. J Virol. 2015;89(1):713-29. doi: 10.1128/JVI.02570-14.
639 PubMed PMID: 25355872; PubMed Central PMCID: PMC4301145.
640 23. Gregory CD, Rowe M, Rickinson AB. Different Epstein-Barr virus-B cell
641 interactions in phenotypically distinct clones of a Burkitt's lymphoma cell line. The
642 Journal of general virology. 1990;71 (Pt 7):1481-95. Epub 1990/07/01. doi:
643 10.1099/0022-1317-71-7-1481. PubMed PMID: 2165133.
644 24. Covarrubias S, Richner JM, Clyde K, Lee YJ, Glaunsinger BA. Host shutoff is a
645 conserved phenotype of gammaherpesvirus infection and is orchestrated exclusively

646 from the cytoplasm. J Virol. 2009;83(18):9554-66. doi: 10.1128/JVI.01051-09. PubMed
647 PMID: 19587049; PubMed Central PMCID: PMC2738246.
648 25. Rowe M, Glaunsinger B, van Leeuwen D, Zuo J, Sweetman D, Ganem D, et al.
649 Host shutoff during productive Epstein-Barr virus infection is mediated by BGLF5 and
650 may contribute to immune evasion. Proceedings of the National Academy of Sciences
651 of the United States of America. 2007;104(9):3366-71. doi: 10.1073/pnas.0611128104.
652 PubMed PMID: 17360652; PubMed Central PMCID: PMC1805610.
653 26. Samusik N, Good Z, Spitzer MH, Davis KL, Nolan GP. Automated mapping of
654 phenotype space with single-cell data. Nature methods. 2016;13(6):493-6. doi:
655 10.1038/nmeth.3863. PubMed PMID: 27183440; PubMed Central PMCID:
656 PMCPMC4896314.
657 27. Maguire O, Collins C, O'Loughlin K, Miecznikowski J, Minderman H. Quantifying
658 nuclear p65 as a parameter for NF-kappaB activation: Correlation between
659 ImageStream cytometry, microscopy, and Western blot. Cytometry A. 2011;79(6):461-9.
660 doi: 10.1002/cyto.a.21068. PubMed PMID: 21520400; PubMed Central PMCID:
661 PMCPMC3140714.
662 28. Rosa MD, Gottlieb E, Lerner MR, Steitz JA. Striking similarities are exhibited by
663 two small Epstein-Barr virus-encoded ribonucleic acids and the adenovirus-associated
664 ribonucleic acids VAI and VAII. Mol Cell Biol. 1981;1(9):785-96. Epub 1981/09/01.
665 PubMed PMID: 9279391; PubMed Central PMCID: PMCPMC369362.
666 29. Virgin HWt, Latreille P, Wamsley P, Hallsworth K, Weck KE, Dal Canto AJ, et al.
667 Complete sequence and genomic analysis of murine gammaherpesvirus 68. J Virol.
668 1997;71(8):5894-904. PubMed PMID: 9223479; PubMed Central PMCID: PMC191845.

669 30. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-
670 PCR. Nucleic Acids Res. 2001;29(9):e45. PubMed PMID: 11328886; PubMed Central
671 PMCID: PMCPMC55695.
672 31. Li Q, He M, Zhou F, Ye F, Gao SJ. Activation of Kaposi's sarcoma-associated
673 herpesvirus (KSHV) by inhibitors of class III histone deacetylases: identification of
674 sirtuin 1 as a regulator of the KSHV life cycle. J Virol. 2014;88(11):6355-67. Epub
675 2014/03/29. doi: 10.1128/JVI.00219-14. PubMed PMID: 24672028; PubMed Central
676 PMCID: PMCPMC4093851.
677 32. Kimball AK, Oko LM, Bullock BL, Nemenoff RA, van Dyk LF, Clambey ET. A
678 Beginner's Guide to Analyzing and Visualizing Mass Cytometry Data. J Immunol.
679 2018;200(1):3-22. Epub 2017/12/20. doi: 10.4049/jimmunol.1701494. PubMed PMID:
680 29255085; PubMed Central PMCID: PMCPMC5765874.
681
682

683 **FIGURE LEGENDS**

684

685 ***Figure 1. RNA-Flow cytometry using the PrimeFlow™ method affords robust and***
686 ***sensitive analysis of endogenous γHV genes at the single-cell level.*** Viral RNA
687 analysis in γHV68-infected fibroblasts by qRT-PCR (A, C) or by flow cytometric analysis
688 using PrimeFlow™ at 16 hpi (B, D, E, F). Samples were either mock, TMER-TKO, or
689 WT γHV68-infected, with infections done using 5 plaque forming units/cell. qRT-PCR
690 standardized to 18s RNA, at the indicated times. All flow cytometric events gated on a
691 generous FSC x SSC gate, followed by singlet discrimination. PrimeFlow™ analysis
692 quantified probe fluorescence for (B) TMERs, (D) gene 73, (E) gene 18 or (F) gene 72
693 relative to side-scatter area (SSC-A). Probe fluorescence is indicated, with all probes
694 detected using either AlexaFluor (AF) 647 or 488 conjugates. Data representative of 2
695 independent experiments, each done with biological replicates.

696

697 ***Figure 2. Heterogeneous gene expression in a γHV68 latently infected B cell line.***
698 Viral, host RNA analysis during γHV68 latency and reactivation in A20.γHV68 (HE2.1)
699 cells by qRT-PCR (A) and flow cytometric analysis using PrimeFlow™ (B-G), comparing
700 untreated or TPA-stimulated samples at 24 hrs post-treatment. Analysis includes A20,
701 virus-negative cells and A20.γHV68 cells. (A) qRT-PCR analysis of TMER-5 expression
702 relative to 18s RNA in A20 and A20.γHV68 (HE2.1) cells, untreated or stimulated with
703 TPA for 24 hrs. (B) PrimeFlow™ detection of TMER expression in A20.γHV68 (HE2.1)
704 cells, comparing either samples that were unstained (solid gray) or stained for the
705 TMERs (open black line). (C) Analysis of TMER expression in multiple conditions,

706 comparing untreated and stimulated A20 and A20.γHV68 cells, with gates defining the
707 frequency of events that expressed either intermediate (mid) or high levels of TMERs.
708 Data depict lymphocytes that were singlets, defined by sequential removal of doublets.
709 (D) Quantification of the frequency of $TMER^{high}$ cells in stimulated A20.γHV68 cells. (E)
710 Histogram overlays of HygroGFP and RCA protein expression in A20.γHV68 cells
711 comparing $TMER^{mid}$ cells from untreated cultures (top, gray), $TMER^{mid}$ cells from TPA-
712 stimulated cultures (middle, blue), with $TMER^{high}$ cells from stimulated cultures (bottom,
713 red). (F) Quantification of the frequencies of HygroGFP+ (left) and RCA protein+ (right)
714 cells as a function of TMER expression and treatment condition. (G) Flow cytometric
715 analysis on an imaging flow cytometer, with each row showing an individual cell and
716 representative images of brightfield (BF), RCA protein (RCA), DAPI, and TMER
717 localization in $TMER^{mid}$ cells from untreated cultures (left) or $TMER^{high}$ RCA+ cells from
718 stimulated (right) A20.γHV68 cells. Data are from two independent experiments, with
719 biological replicates within each experiment for all A20.γHV68 cultures. Graphs depict
720 the mean ± SEM, with each symbol identifying data from a single replicate. Statistical
721 analysis was done using an unpaired t test (D) or one-way ANOVA, subjected to
722 Tukey's multiple comparison test (A, F), with statistically significant differences as
723 indicated, * $p<0.05$, **** $p<0.0001$.
724
725 ***Figure 3. Single-cell analysis of KSHV PAN RNA expression in the BCBL-1 B cell***
726 ***lymphoma cell line.*** (A) Flow cytometric analysis of PAN RNA expression in multiple
727 conditions, from cells incubated with no probe (left), or cells subjected to hybridization
728 using a probe for PAN RNA, comparing virus-negative BL41 cells (second from left) with

729 KSHV+ BCBL-1 cells that were either untreated or stimulated (with TPA and sodium
730 butyrate (NaB)) for 72 hours. Data depict lymphocytes that were singlets, defined by
731 sequential removal of doublets. Representative images were defined as samples that
732 were closest to the median frequency. Quantification of (B) the frequency of PAN RNA+
733 cells and (C) PAN RNA median fluorescence within PAN RNA+ cells, comparing
734 untreated or stimulated BCBL-1 cells. D, E) Flow cytometric analysis of untreated
735 BCBL-1 cells, using histogram overlays, to compare cell size (forward scatter, FSC),
736 granularity (side scatter, SSC) and gene 73 expression in cells that were either PAN
737 RNA negative [-] (blue line) or PAN RNA positive [+] (red line). Data show (D) histogram
738 overlays of these populations with fluorescence quantification provided in panel E. Gene
739 73 analysis includes samples in which there was no gene 73 probe (i.e. “No Probe”, in
740 solid gray), to define background fluorescence. (F, G) Flow cytometric analysis of
741 stimulated BCBL-1 cells, using histogram overlays, to compare cell size (forward
742 scatter, FSC), granularity (side scatter, SSC) and gene 73 expression in cells that are
743 either PAN RNA negative [-] (blue) or PAN RNA positive [+] (red). Data show (F)
744 histogram overlays of these populations with fluorescence quantification provided in
745 panel G, with gene 73 analysis including a “No Probe” sample (gray) to define
746 background fluorescence. Due to variable baseline fluorescence values for SSC and
747 gene 73 between experiments, values were internally standardized to fluorescent
748 intensities within the PAN RNA negative population for each experiment, with data
749 depicting mean ± SEM. Symbols in panels B and C indicate values from individual
750 samples. Data are from two independent experiments, with biological replicates within
751 each experiment, with total number of biological replicates as follows: No Probe (n = 2),

752 BL41 control (n = 3), BCBL-1 untreated (n = 6), BCBL-1 stimulated (n = 6). Statistical
753 analysis was done using an unpaired t test with statistically significant differences as
754 indicated, * $p<0.05$, ** $p<0.01$, **** $p<0.0001$.
755
756 ***Figure 4. Single-cell analysis of EBV EBER expression in the Mutu I B cell***
757 ***lymphoma cell line.*** (A) Flow cytometric analysis using the PrimeFlow^{TM} method, to
758 quantify EBER expression in multiple conditions, from cells incubated with no probe
759 (left), or cells subjected to hybridization using a probe for the EBERs, comparing virus-
760 negative BL41 cells (second from left) with Mutu I EBV+ cells that were either untreated
761 (with DMSO) or stimulated (with TPA in DMSO) for 48 hours. Data depict lymphocytes
762 that were singlets, defined by sequential removal of doublets. Representative images
763 were defined as samples that were closest to the median frequency, with data depicting
764 mean ± SEM. (B) Quantitation of median EBER fluorescence within EBER+ Mutu I cells
765 in untreated and stimulated cultures. Horizontal dashed line indicates the background
766 fluorescent signal from BL41 controls. Data are from two independent experiments, with
767 biological replicates within each experiment, with total biological replicates as follows:
768 No Probe (n = 2), BL41 Control untreated (n = 6), Mutu I untreated (n = 6), Mutu I
769 stimulated (n = 6). Statistical analysis was done using an unpaired t test with statistically
770 significant differences as indicated, * $p<0.05$.
771
772 ***Figure 5. Actin mRNA degradation identifies virally-infected cells experiencing***
773 ***virus-induced host shutoff.*** Actin mRNA analysis by qRT-PCR (A,D) or by flow
774 cytometric analysis using PrimeFlow^{TM} (B,C,E-G), comparing cells with variable

775 infection status. (A) qRT-PCR analysis of beta-actin (Actb) mRNA expression relative to
776 18s RNA in mock, WT or TMER-TKO infected 3T12 fibroblasts at 18 hpi. (B)
777 PrimeFlow^{TM} analysis of actin mRNA in 3T12 fibroblasts, either unstained ("No probe"),
778 mock-infected or infected with WT γHV68 or TMER-TKO at 18 hpi. (C) PrimeFlow^{TM}
779 analysis of actin mRNA and TMER expression in WT γHV68 infected fibroblasts at 18
780 hpi. (D) qRT-PCR analysis of beta-actin (Actb) mRNA expression relative to 18s RNA
781 in A20, virus-negative cells and A20.γHV68 (HE2.1) cells, untreated or stimulated with
782 TPA for 24 hrs. (E) PrimeFlow^{TM} analysis of actin mRNA and TMER expression in
783 untreated and stimulated A20.γHV68 cells, with the frequency of TMER^{high} actin
784 mRNA^{low} cells indicated, based on the gated events. (F,G) Actin mRNA analysis by
785 PrimeFlow^{TM} using either (F) histogram overlays or (G) quantifying frequencies,
786 comparing A20.γHV68 cells that were either untreated or stimulated, further stratified by
787 whether the cells were TMER^{mid} or TMER^{high} (using the gating strategy defined in Fig.
788 2C). All flow cytometry data depict single cells, defined by sequential removal of
789 doublets. Data are from two-three independent experiments, with biological replicates
790 within each experiment. Graphs depict the mean ± SEM, with each symbol identifying
791 data from a single replicate. Statistical analysis was done using one-way ANOVA,
792 subjected to Tukey's multiple comparison test (A, D, G), with statistically significant
793 differences as indicated, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$.
794
795 ***Figure 6. Heterogeneous viral gene expression at the single-cell level during lytic***
796 ***replication.*** Viral, host RNA flow cytometric analysis in γHV68-infected fibroblasts at 16
797 hpi defined by the PrimeFlow^{TM} method, comparing (A) X-shift clustering analysis and

798 (B-E) biaxial gating analysis for the indicated features. (A) Automated, clustering
799 analysis using the X-Shift algorithm on 10,000 events total, compiled from mock- and
800 γHV68-infected fibroblasts at 16 hrs pi (1,000 events randomly imported per sample,
801 mock infected n=4, γHV68-infected n=6) identifies multiple clusters of cells with
802 differential gene expression (7 clusters, colored distinctly, "Cluster ID"), with these
803 clusters then depicted for expression of TMERs, gene 73, Actin mRNA, and RCA.
804 Range of expression is identified for each parameter. (B) Analysis of TMER and gene
805 73 co-expression in mock (left), WT γHV68-infected (middle), and TMER-TKO-infected
806 (right) samples, with gates depicting populations with different gene expression profiles,
807 defined relative to mock and TMER-TKO infected samples. (C) Color-coded populations
808 from WT-infected sample in panel B, with each color indicating a different gene
809 expression profile. (D) Histogram overlays of the five populations identified in panel C
810 for the indicated parameters. (E) Quantitation of gene expression among the five
811 populations identified in panel C, using the same color-coding strategy. Data are from
812 three independent experiments, with each experiment containing biological replicates.
813 Flow cytometry data shows single cells that are DNA+ (DAPI+). Statistical significance
814 tested by one-way ANOVA, comparing the mean of $TMER^{high}$ gene 73^{high} cells to all
815 other means, followed by Dunnett's multiple testing correction. Significance identified as
816 *** $p<0.001$, **** $p<0.0001$.
817
818 ***Figure 7. Lytic replication is characterized by heterogeneous TMER localization***
819 ***and variable penetrance of actin RNA degradation.*** Viral, host RNA flow cytometric
820 analysis in γHV68-infected fibroblasts at 16 hpi defined by the PrimeFlow™ method. (A)

821 The frequency of γHV68-infected fibroblasts with TMERs primarily in the nucleus was
822 quantified by ImageStream, with data showing the frequency of cells in which
823 TMER:DAPI colocalization (i.e. similarity score) was >1. (B) Images showing brightfield
824 (BF), TMER, RCA protein (RCA), DAPI, gene 73 and actin mRNA localization,
825 comparing cells with nuclear TMER localization (left) versus cytoplasmic TMER
826 localization (right). (C) Analysis of cell subpopulations stratified by TMER and gene 73
827 expression (defined in Fig. 6C), subjected to the tSNE dimensionality reduction
828 algorithm. Data and cell populations are derived from the dataset presented in Fig. 6,
829 showing all DNA+ (DAPI+) single cells (FSC-A, SSC-A) subjected to the tSNE
830 algorithm. The tSNE algorithm provides each cell with a unique coordinate according to
831 its expression of Actin mRNA, RCA, Gene 73, and TMERs, displayed on a two-
832 dimensional plot (tSNE1 versus tSNE2). Visualization grid of tSNE plots, with plots
833 arranged according to marker expression (rows) relative to phenotype of the cellular
834 population examined (columns). (D) Biaxial analysis of actin RNA versus RCA protein,
835 among the five populations identified in Fig. 6. Flow cytometry data shows single cells
836 that are DNA+ (DAPI+). Data in panels A-B from two independent experiments, in panel
837 C-D from three independent experiments.

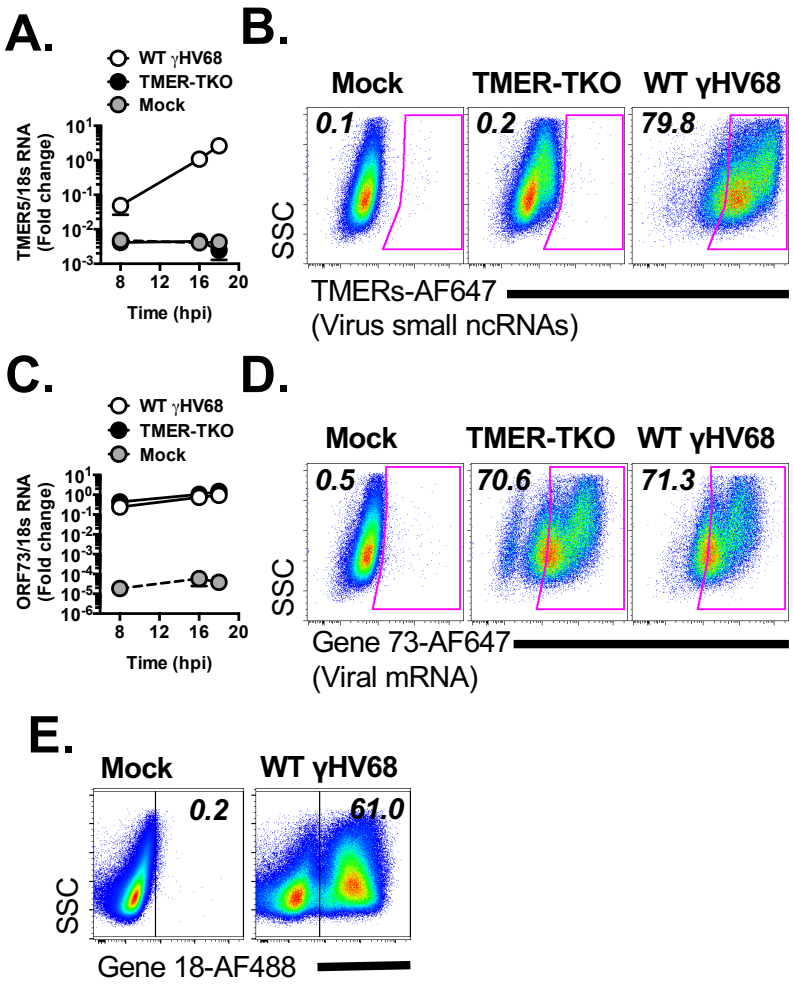

Figure 1
Oko et al

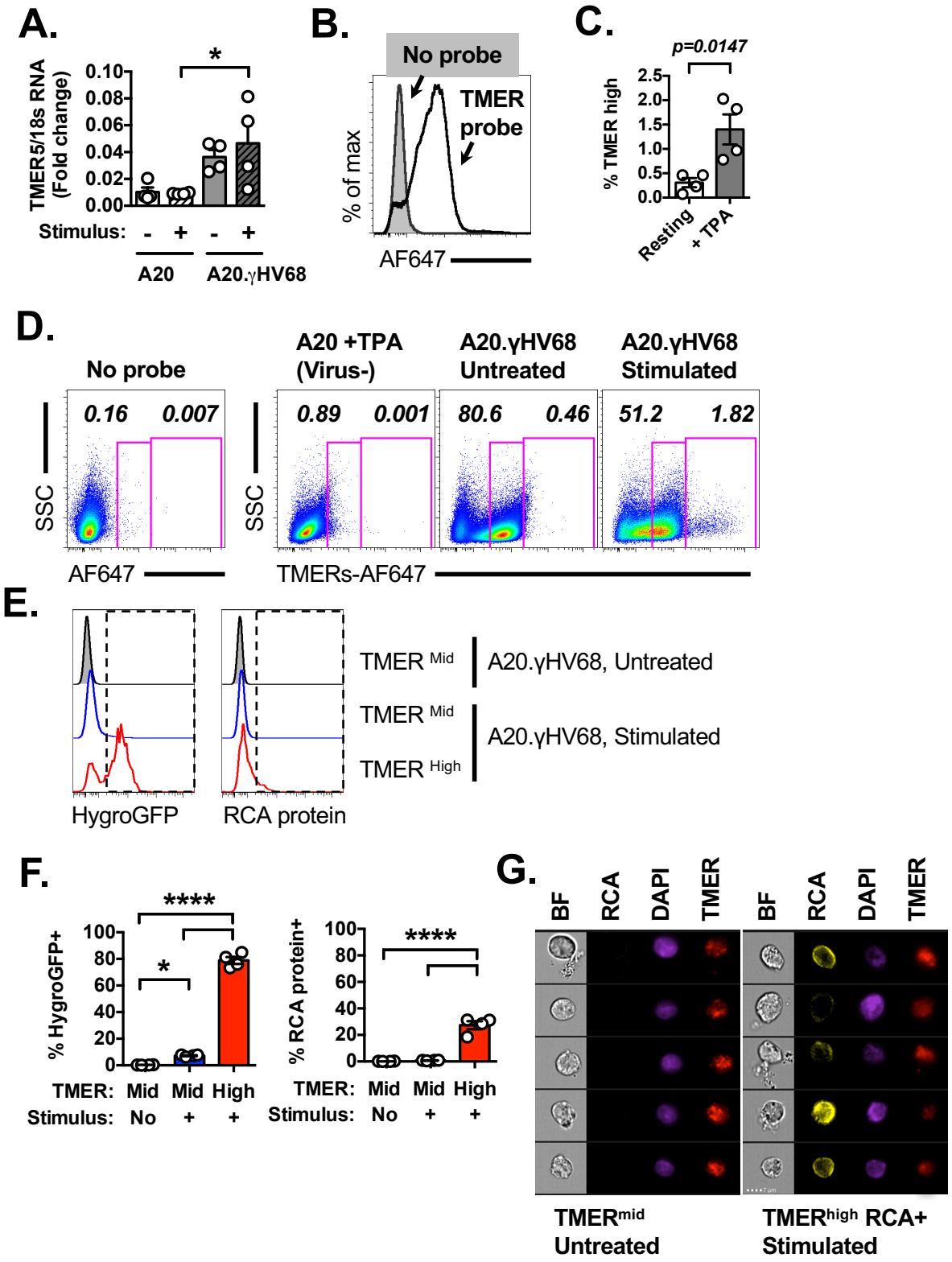

Figure 2
Oko et al

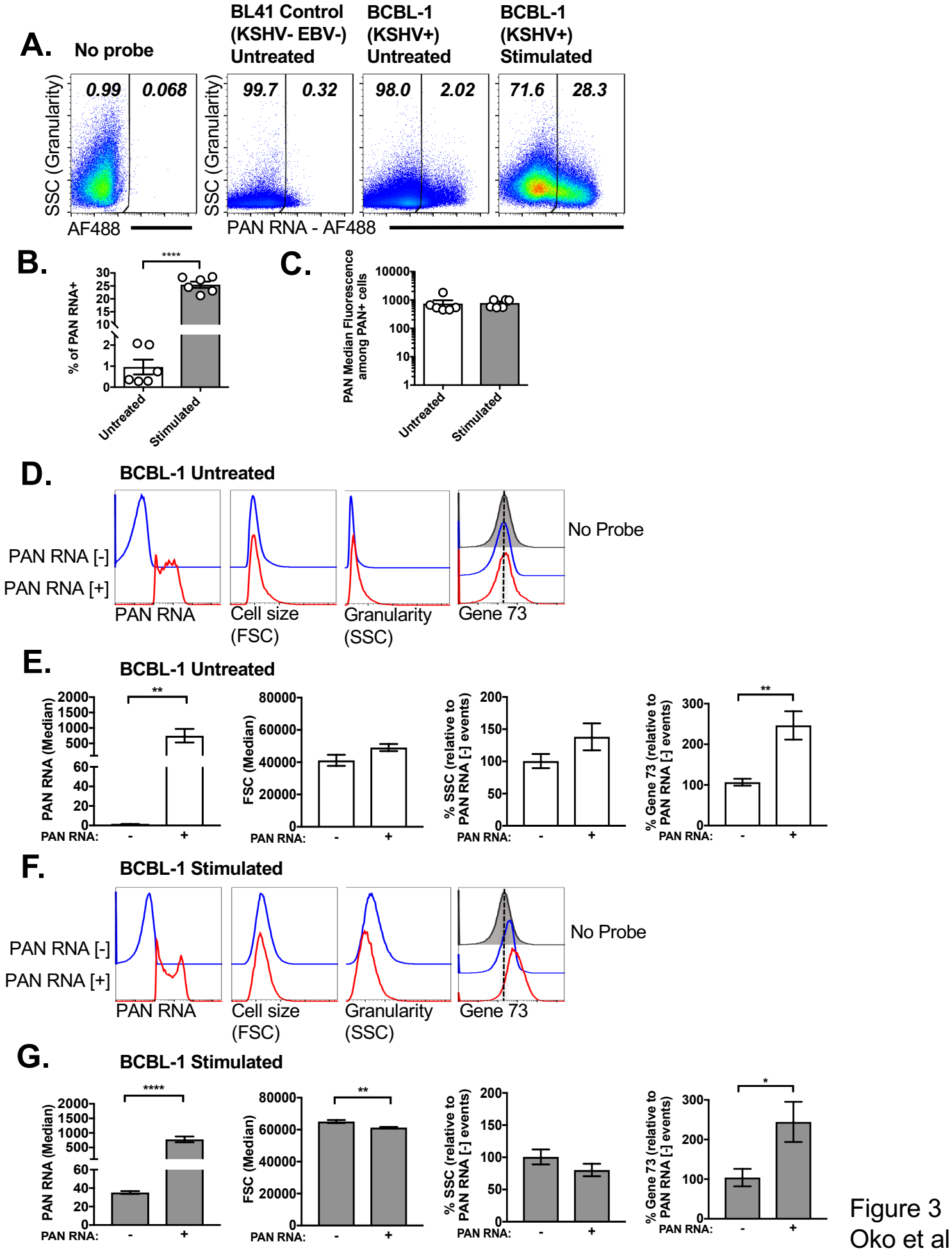

Figure 3
Oko et al

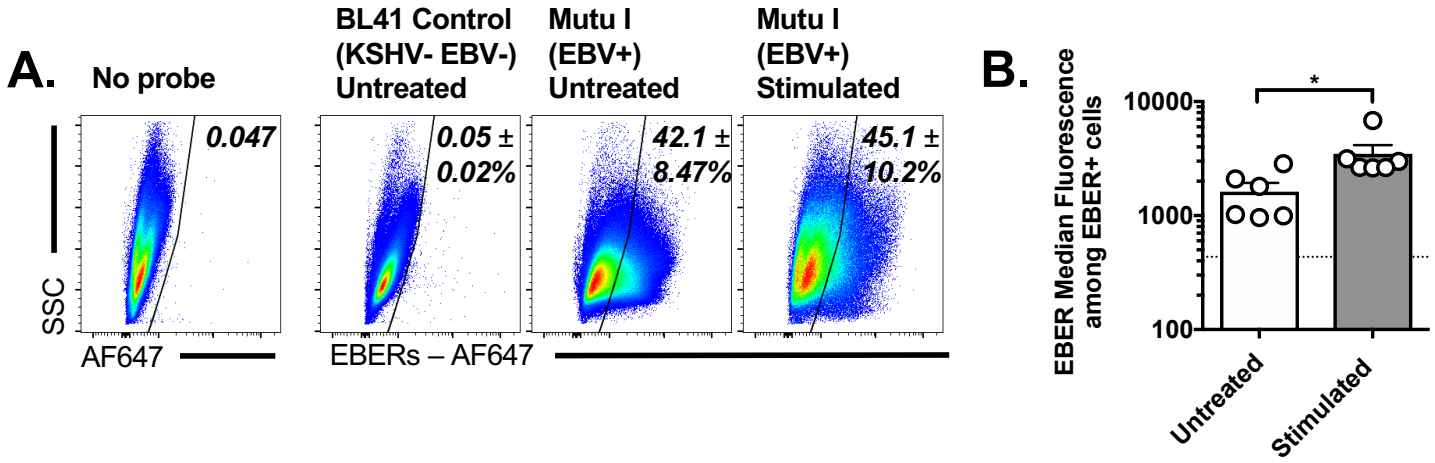

Figure 4
Oko et al

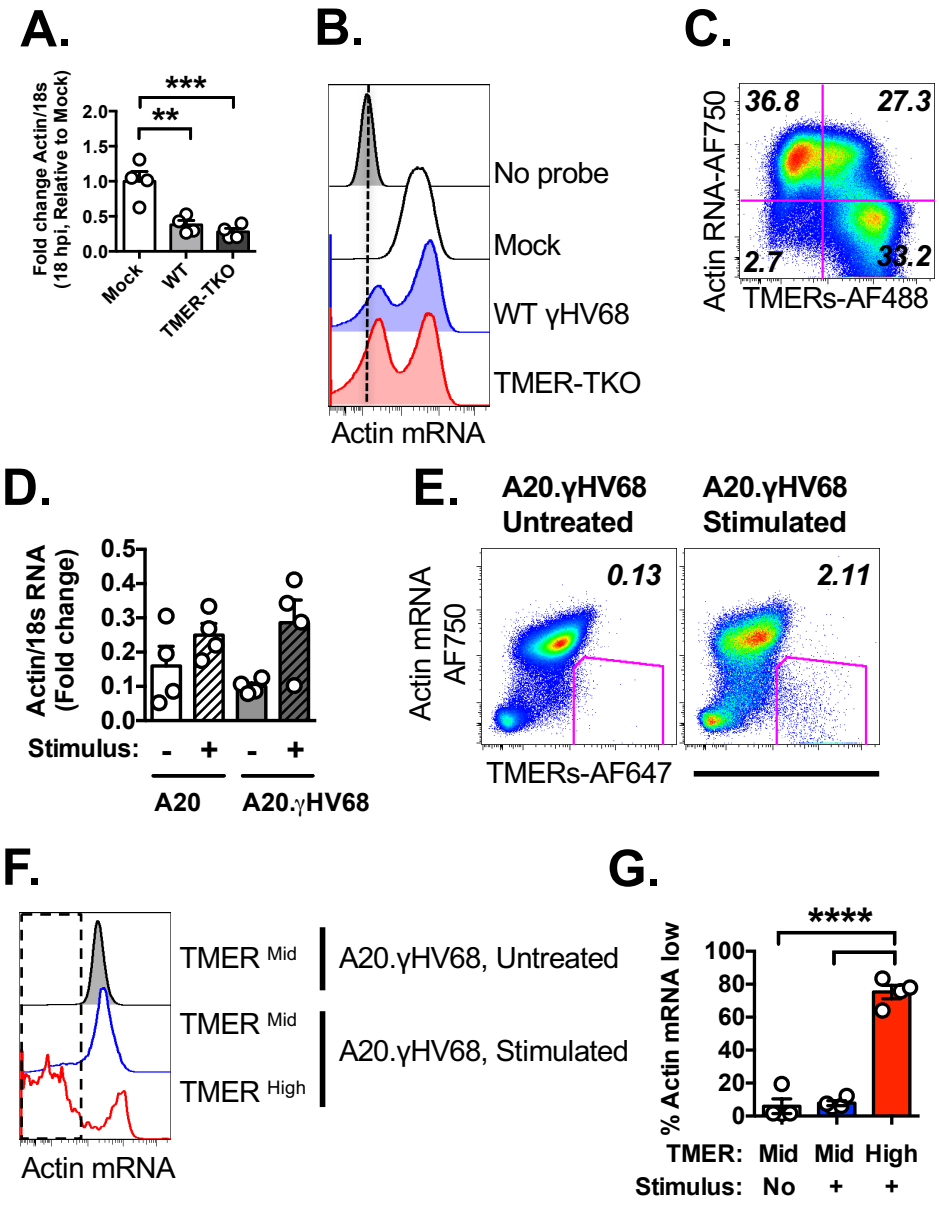

Figure 5
Oko et al

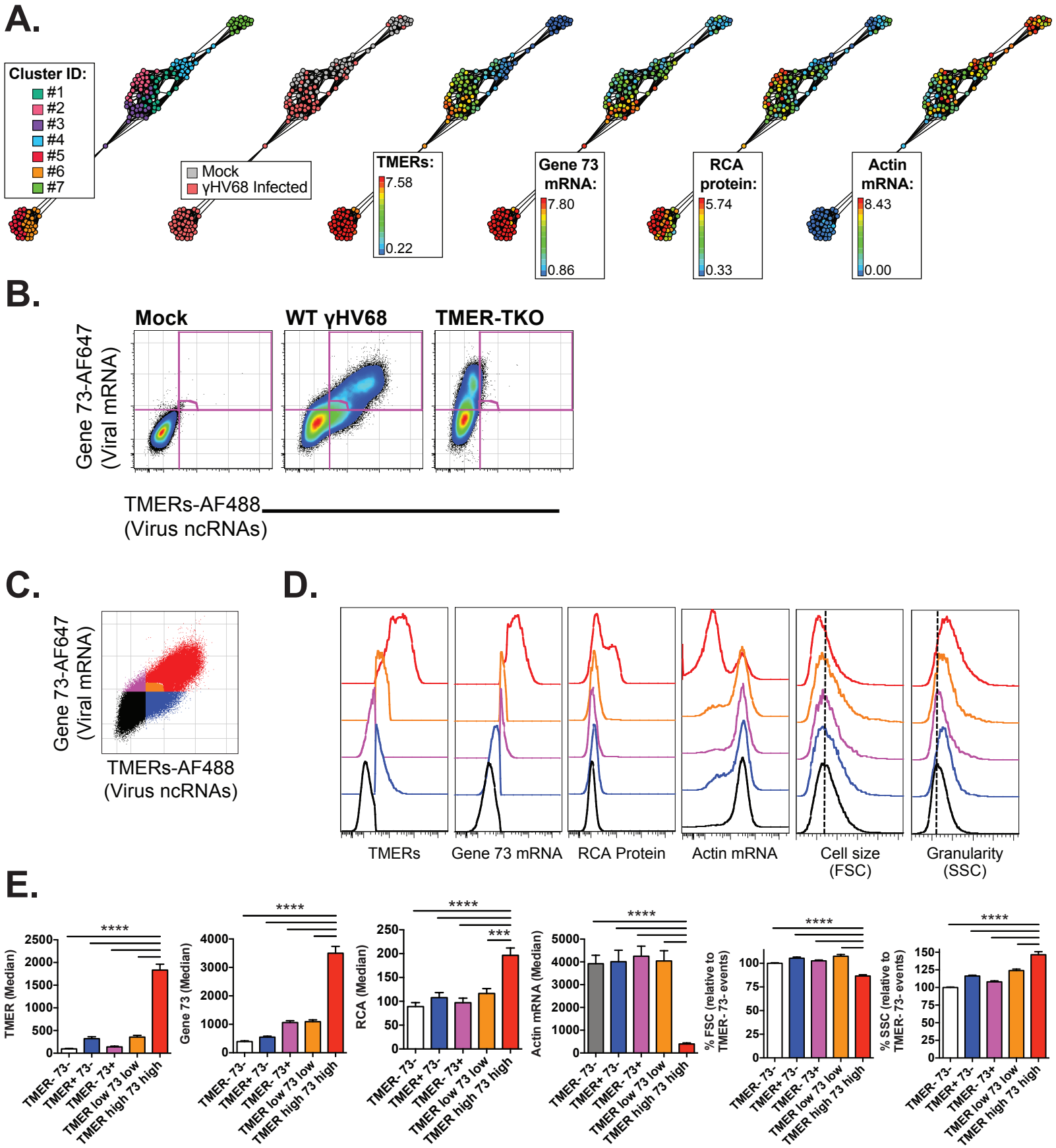

Figure 6
Oko et al

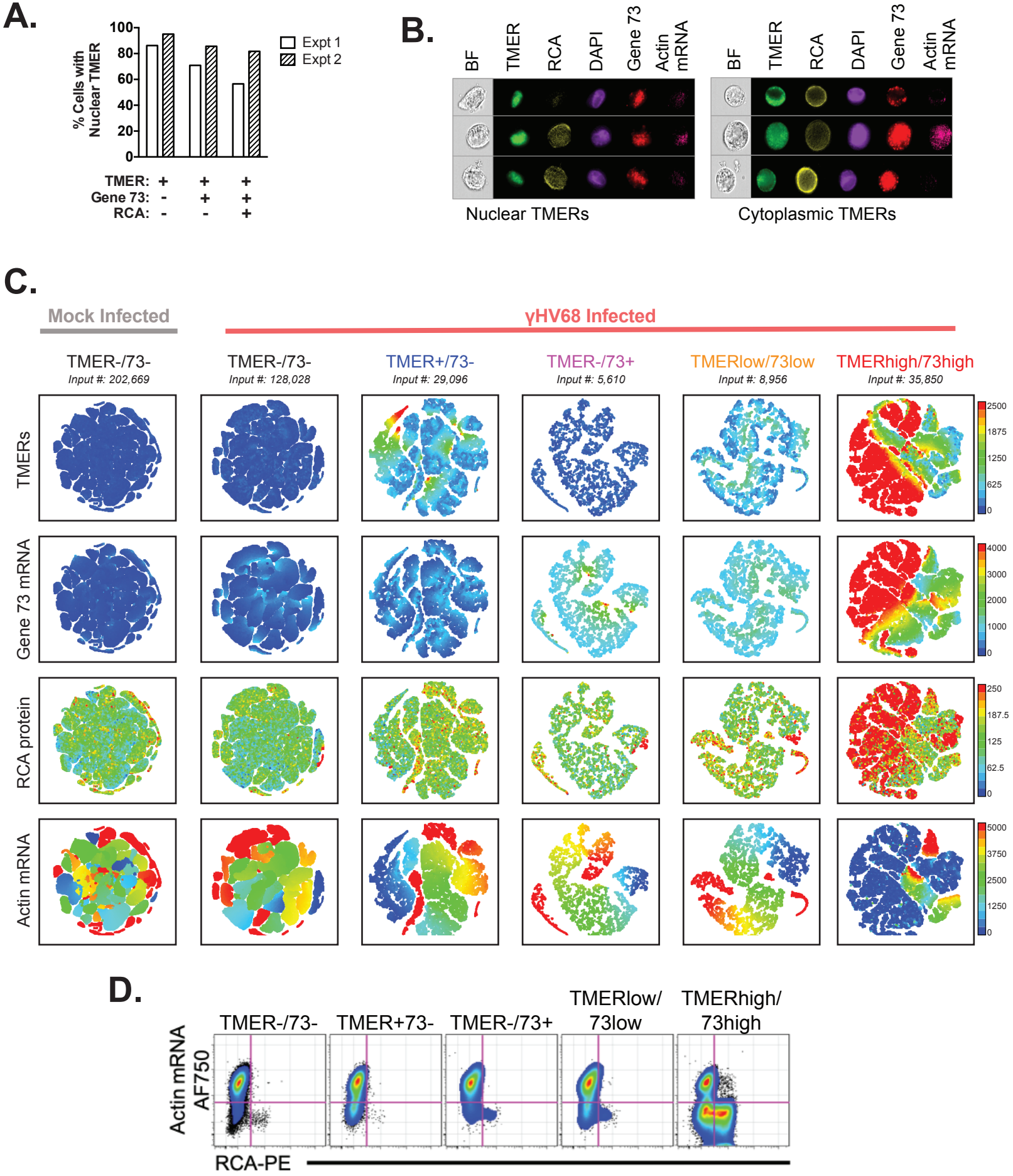

Figure 7
Oko et al