Fate-resolved gene regulatory signatures of individual B lymphocytes in the early stages of Epstein-Barr Virus infection

Elliott D. SoRelle[1,2*], Joanne Dai[1,3], Nicolás M. Reinoso-Vizcaino[1], Ashley P. Barry[1], Cliburn Chan[2], & Micah A. Luftig[1*]

[1] Department of Molecular Genetics and Microbiology, Duke Center for Virology, Duke University School of Medicine, Durham, NC 27710
[2] Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, NC 27710
[3] Current address: Amgen Inc., 1120 Veterans Blvd, South San Francisco, CA 94080

*Corresponding Authors: Elliott D. SoRelle & Micah A. Luftig

Abstract

Epstein-Barr Virus (EBV) infection of B lymphocytes elicits diverse host responses via complex, well-adapted transcriptional control dynamics. Consequently, this host-pathogen interaction provides a powerful system to explore fundamental cellular processes that contribute to consensus fate decisions including cell cycle arrest, apoptosis, proliferation, and differentiation. Here we capture these responses and fates with matched single-cell transcriptomics and chromatin accessibility, from which we construct a genome-wide multistate model of early infection dynamics. Notably, our model captures a previously uncharacterized EBV$^+$ analog of a multipotent activated precursor state that can yield early memory B cells. We also find that a marked global reduction in host chromatin accessibility occurs during the first stages of infection in subpopulations of EBV$^+$ cells that display senescent and pre-apoptotic hallmarks induced by innate antiviral sensing and proliferation-linked DNA damage. However, cells in proliferative infection trajectories exhibit greater accessibility at select host sites linked to B cell activation and survival genes as well as key regions within the viral genome. To further investigate such loci, we implement a bioinformatic workflow (crisp-ATAC) to identify phenotype-resolved regulatory signatures. This customizable method applies user-specified logical criteria to produce genome-wide single-cell ATAC- and ChIP-seq range intersections that are used as inputs for *cis*-linkage prediction and ontology tools. The resulting tri-modal data yield exquisitely detailed hierarchical perspectives of the transforming regulatory landscape during critical stages of an oncogenic viral infection that simulates antigen-induced B cell activation and differentiation. We anticipate these resources will guide investigations of gene regulatory modules controlling EBV-host dynamics, B cell effector fates, and lymphomagenesis.

38 **Introduction**

39

40 Epstein-Barr Virus (EBV) is an oncogenic gammaherpesvirus present in >90% of adults
41 (Rickinson and Kieff, 2007) and associated with up to 2% of human cancers (Cohen et al., 2011).
42 Recent reports have also provided epidemiological and mechanistic evidence supporting an
43 etiological role for EBV in Multiple Sclerosis (MS) (Bjornevik et al., 2022; Lanz et al., 2022). In its
44 initial stages, EBV infection within primary B lymphocytes manifests an array of host and viral
45 programs. Upon entry into the host cell, the linear dsDNA viral genome rapidly circularizes to form
46 an episome that is retained within the nucleus (Lindahl et al., 1976; Nonoyama and Pagano,
47 1972). Within hours to days, host innate immune responses are generated to restrict viral
48 progression (Lünemann et al., 2015; Martin et al., 2007; Smith et al., 2013; Tsai et al., 2011).
49 Simultaneously, viral genes are expressed to counteract host defenses (Ressing et al., 2015), co-
50 opt B cell-intrinsic activation and proliferation (Calender et al., 1987; Thorley-Lawson, 2001;
51 Thorley-Lawson and Mann, 1985), and attenuate DNA damage and stress responses instigated
52 by virus-induced growth (McFadden et al., 2016; Nikitin et al., 2010). A consequence of these
53 intimately adapted host-pathogen dynamics is that EBV infection can precipitate diverse
54 responses and outcomes for host B cells. These include unsuccessful infection routes resulting
55 from effective antiviral restriction and DNA damage-induced growth arrest as well as successful
56 infection leading to immortalization *in vitro* (Bird, 1981; Henle et al., 1967; Pope et al., 1968; Zhao
57 et al., 2011) or lifelong latency *in vivo* within memory B cells (Babcock, 1998; Longnecker et al.,
58 2013; Miyashita et al., 1997) that retain oncogenic potential (Raab-Traub, 2007; Thorley-Lawson
59 and Gross, 2004).

60 Since its discovery in 1964 as the first human tumor virus (Epstein et al., 1964; Young and
61 Rickinson, 2004), extensive research has revealed the molecular means by which EBV
62 establishes infection and underlies various malignancies. The entire EBV genome is ~172
63 kilobases and contains at least 80 protein-coding sequences including six EBV nuclear antigens
64 (EBNAs); several latent membrane proteins (LMPs); and loci that encode replicative and
65 transcriptional machinery as well as structural proteins. The EBV genome also contains functional
66 non-coding RNAs: the BHRF and BART microRNAs and the EBV-encoding regions (EBERs)
67 (Rickinson and Kieff, 2007; Young et al., 2007).

68 EBNAs are especially important in establishing distinct forms of latency depending on their
69 combinatorial expression (Price and Luftig, 2015). EBNA1 is a transcription factor (TF) that is
70 essential for viral genome maintenance and B cell transformation and ubiquitously binds and
71 epigenetically regulates host chromatin (Altmann et al., 2006; Canaan et al., 2009; Dheekollu et

72 al., 2021; Humme et al., 2003; Lamontagne et al., 2021; Lu et al., 2010; Lupton and Levine, 1985;
73 Wood et al., 2007; Yates et al., 1985). EBNALP is another essential factor (Mannick et al., 1991;
74 Szymula et al., 2018) that initiates host cell proliferation alongside its co-activated target, EBNA2
75 (Alfieri et al., 1991; Harada and Kieff, 1997; Sinclair et al., 1994), and interacts with several host
76 proteins including TFs (Han et al., 2001; Ling et al., 2005; Matsuda et al., 2003). EBNA2 is likewise
77 required for B cell immortalization (Cohen et al., 1989), notably through coordination with host
78 TFs and their binding sites (Lu et al., 2016; Zhao et al., 2011) and with EBNALP to drive early cell
79 proliferation and viral LMP1 expression (Peng et al., 2005). The EBNA3 proteins (EBNA3A,
80 EBNA3B, and EBNA3C) mediate a delicate balance of anti- and pro-oncogenic processes (Allday
81 et al., 2015; Banerjee et al., 2014; Parker et al., 1996; Tomkinson et al., 1993; White et al., 2010).
82 These include epigenetic repression of host tumor suppressor genes (*BIM, p14, p16*) and viral
83 promoters (Maruo et al., 2011; Paschos et al., 2012; Saha et al., 2015; Skalska et al., 2010; Styles
84 et al., 2017), competitive binding of the EBNA2-interacting host factor RBPJ (Robertson et al.,
85 1995; Wang et al., 2015), and inhibition of apoptosis (Price et al., 2017). Collectively, the EBNAs
86 reshape the nuclear regulatory and transcriptional landscape of EBV$^+$ B cells, effectively hijacking
87 B cell-intrinsic activation, expansion, and differentiation programs. Thus, EBV co-opts antigen-
88 responsive host immune mechanisms for the ulterior purposes of viral replication and
89 propagation.

90 While the EBNAs engage cell proliferation machinery at the epigenetic and transcriptional level
91 in the nucleus, the LMPs (LMP1, LMP2A, and LMP2B) do so at the cell membrane by simulating
92 antigen-induced signal transduction pathways. The essential LMP1 promotes B cell activation
93 through mimicry of a constitutively active CD40 receptor (Kilger et al., 1998; Uchida et al., 1999)
94 and interacts with Tumor Necrosis Factor (TNF) receptor-associated factors (TRAFs) to activate
95 NF-κB pathway signaling via IKK (Devergne et al., 1996; Eliopoulos et al., 2003; Greenfeld et al.,
96 2015; Luftig et al., 2003). These interactions induce anti-apoptotic pathways, MHC-mediated
97 immune recognition, pro-inflammatory responses, and cell migration. Downstream consequences
98 include oncogenic proliferation and survival but also induction of pro-apoptotic responses
99 (Devergne et al., 1998; Fries et al., 1999; Greenfeld et al., 2015; Henderson et al., 1991; Shair et
100 al., 2008; Wang et al., 2017). Thus, as in antigen-induced B cell activation (and subsequent
101 differentiation), adept regulatory control of NF-κB signaling (Hoffmann et al., 2002; Mitchell et al.,
102 2018; O'Dea et al., 2007; Roy et al., 2019) is dispositive for the fate of a given EBV$^+$ B cell.
103 Although it is not essential for transformation, LMP2A promotes cell survival through mimicry of a
104 stimulated B cell receptor (BCR), which activates signaling cascades complementary to those
105 induced by LMP1 (Anderson and Longnecker, 2008; Fish et al., 2020; Guasparri et al., 2008;

106 Portis and Longnecker, 2004). LMP2A expression further predisposes EBV$^+$ B cells to survival by
107 lowering antigen affinity selection thresholds *in vivo* (Minamitani et al., 2015). Thus, EBV latent
108 membrane proteins play integral roles in B cell proliferation in the absence of antigen licensing
109 and in avoiding replicative dead ends effected by antiviral sensing.
110 Clearly, key EBV gene products manipulate diverse host programs at early stages to achieve
111 sustained latency (Mrozek-Gorska et al., 2019; Pich et al., 2019). Many such perturbations involve
112 extensive rewiring of epigenetic and transcriptional regulatory modules. EBV researchers have
113 used methods such as RNA-, ATAC- (Assay for Transposase-Accessible Chromatin), and ChIP-
114 seq (Chromatin Immunoprecipitation) to study these changes at various levels in the gene
115 regulatory hierarchy within early infected cells and transformed lymphoblastoid cell lines (LCLs)
116 (Arvey et al., 2012; Jiang et al., 2017; McClellan et al., 2013; Mrozek-Gorska et al., 2019; Wang
117 et al., 2019; Zhou et al., 2015). Recently, the epigenetic and transcriptional roles of EBNA1 were
118 interrogated through time-resolved multi-omics (Lamontagne et al., 2021). While these and other
119 studies provide indispensable insights regarding virus-induced genome-wide expression and
120 regulation, they have relied on bulk ensemble sequencing. Such assays yield population-
121 averaged measurements that obscure variation arising from intrinsic stochasticity (Raj et al.,
122 2006; Raj et al., 2010; Raj and Van Oudenaarden, 2008), asynchronous behaviors, and
123 heterogeneous cell subsets. Specifically, ensemble averaging fails to capture cell-matched
124 measurements across genes, which precludes identification of coordinated expression programs
125 or epigenomic regulatory patterns in specific phenotypes. By contrast, single-cell sequencing
126 provides refined genome-wide views of expression and regulation that preserve the ability for the
127 identification of heterogenous cell states with low bias (Buenrostro et al., 2015; Junker and van
128 Oudenaarden, 2014; Shalek et al., 2013; Shapiro et al., 2013; Wills et al., 2013). Given the
129 complexity of host-virus relationships, single-cell -omics approaches are essential to dissect the
130 early stages of EBV infection and the distinct fate trajectories it comprises. We previously used
131 single-cell RNA sequencing (scRNA-seq) to identify EBV-driven heterogeneity in LCLs (SoRelle
132 et al., 2021). Recent advances in single-cell multimodal -omics methods have made it possible to
133 integrate scRNA-seq with several levels of hierarchical regulation (Efremova and Teichmann,
134 2020), which can provide greater insight into the mechanistic origins underlying gene expression.
135 These include techniques for obtaining cell-matched measurements of mRNA transcripts,
136 chromatin accessibility, and DNA methylation status (Cao et al., 2018; Chen et al., 2019; Clark et
137 al., 2018; Zhu et al., 2019), as well as other molecular levels. In this work, we leverage single-cell
138 multiomics (scRNA-seq + scATAC-seq) to capture and explore the distinct gene expression and
139 regulatory signatures that determine the course of EBV infection in primary human B lymphocytes.

Results

EBV asynchronously induces primary B cells into distinct phenotypic states early after infection

To interrogate chromatin accessibility and gene expression changes that occur upon EBV infection, we isolated primary human B cells from the peripheral blood of two donors and infected them with the B95-8 strain of EBV. Infections were performed at a multiplicity of infection (MOI) of 5 to ensure latent gene expression in every cell (Nikitin et al., 2010). We cryopreserved samples of infected cells at 2-, 5-, and 8-days post-infection in addition to uninfected cells (Day 0) from each donor sample following B cell enrichment. Cell samples from each donor and timepoint were simultaneously thawed, prepared to >90% viability, and processed into single-cell multiome libraries. Single-cell matched transcript and accessibility data were obtained through standard NGS, alignment, counting, and quality control (QC) methods (**Table S1**).

EBV infection induced broad transcriptomic changes in B lymphocytes at high efficiency, as evidenced by the near-complete loss of resting phenotypes (Day 0) within two days of infection. New states emerged between Day 2 and Day 5, while subtle shifts in state proportions defined the period between Day 5 and Day 8 (**Figure 1A**). Total and unique transcripts per cell increased, particularly between Day 0 and Day 2, while mitochondrial gene expression increased gradually (**Figure 1B**). Total transcript and mitochondrial distributions at Day 2 exhibited two modes, which was consistent with the presence of both non-proliferative and mitotic cells identified by S-phase and G2M-phase marker scoring (**Figure 1C**).

Unsupervised methods revealed subpopulations (clusters) in cell cycle-regressed aggregated scRNA-seq time courses. Two clusters corresponded to uninfected B cells (c3, c8); seven were post-infection B cell phenotypes (c0, c1, c2, c4, c5, c6, c7); and two were T cells (c9) and $CD14^+$ monocytes (c10) carried over from PBMCs despite extensive B-cell enrichment (**Figure 1D**). Genome-wide expression correlation was higher among post-infection states relative to uninfected cells, and certain phenotypes were more strongly correlated (e.g., c0 with c1; c4 with c7, **Figure 1E**). Sorting cluster membership by day yielded coarse-grained dynamics of cell state transitions (**Figure 1F**). We determined top differential genes in each cluster (one-vs-all-others) to inform state identity annotations (**Figure 1G**). Identified clusters included many genes known to be modulated in EBV infection and were broadly consistent across both donors with respect to top marker genes, cell population frequencies, and temporal emergence (**Figures S1-S4**).

174 Infected cell state heterogeneity is linked to antiviral and B cell-intrinsic responses

175 Cluster analysis deconvolved heterogeneous biological states within each sample and
176 revealed phenotypes retained across multiple timepoints (**Figure 2**). At this resolution, we
177 estimated time- and state-level trends in viral gene expression, variation in metabolic activity, and
178 transcript diversity (**Figure 2A**). Overall, c0, c1, c2, c5, and c6 exhibited the highest levels of EBV
179 transcripts and more unique transcripts than c3, c4, c7, and c8. Mitochondrial gene fraction and
180 unique feature content were highest in c6 and lowest in c4 and c7, although c7 had a long-tailed
181 distribution of mitochondrial expression (20-80%) prior to QC, indicative of (pre-) apoptotic cells.
182 All clusters except uninfected B cells (c3, c8) displayed broad innate antiviral and interferon-
183 stimulated gene (*ISG* and *IFI* member) expression. Antiviral gene expression was generally higher
184 and exhibited greater variance in c7 than c4 and persisted at roughly uniform levels in c0, c1, c2,
185 c5, and c6 (**Figure 2B**). Through joint consideration of cluster-resolved expression trends for viral,
186 mitochondrial, and interferon-stimulated genes, we distinguished uninfected cells (c3, c8) and two
187 classes of cells with the hallmarks of antiviral response: those with low proliferation and negligible
188 viral expression (c4, c7) and those with viral and metabolic indicators of progressive EBV infection
189 (c0, c1, c2, c5, c6).

190 Next, we extensively analyzed differentially expressed genes (DEGs) among clusters and
191 groups, including pairwise comparisons of all post-infection phenotypes (**Figures 2C, S5-S11**).
192 The two resting cell phenotypes differed in their expression of *IGHD*, *IGHM*, *CD27*, and other
193 markers that distinguish naïve (c8) from memory (c3) B cells. In addition to interferon response
194 signatures, non-proliferating infected cells exhibited an overall reduction in gene expression and
195 upregulated stress response markers. These included the highest overall expression of actin
196 sequestration genes (*TMSB10*, *TMSB4X*) and, particularly within c7, numerous ribosomal subunit
197 genes (e.g., *RPS27A*). Cells in c4 were distinguished by elevated expression of *MARCH1*, which
198 encodes an E3 ubiquitin ligase that regulates the type I interferon response (Wu et al., 2020).
199 Unlike c4, c7 cells also contained high transcript levels for genes involved in oxidative stress
200 (*TXN*, *FTL*, *FTH1*), cytochrome oxidase subunits (e.g., *COX7C*), ubiquitin genes (*UBA52*, *UBL5*)
201 and highly variable mitochondrial fractions. Among EBV$^+$ cells with hallmarks of elevated
202 respiration, those in c6 were most clearly consistent with proliferating cells based on upregulated
203 cell cycle markers. Cells in c0 were distinguished by upregulation of *FCRL5* and *LY86-AS1*, an
204 antisense RNA to a lymphocyte antigen (*LY86*) that mediates innate immune responses. Cells in
205 this cluster also displayed markers consistent with the early stages of pre- germinal center
206 activated B cells (e.g., *CCR6, CD69, POU2AF1, TNFRSF13B, PIK3AP1*). Notably, cells in c1 and
207 c2 contained the highest levels of the EBV gene *BHRF1*. Between these two phenotypes, c2 was

208 enriched for genes involved in NF-κB signaling and known markers of EBV-mediated B cell
209 activation (*NFKBIA*, *TNFAIP3*, *EBI3*) while c1 appeared to be derived from naïve cells (based on
210 *IGHD* and other carryover genes) and exhibited near-unique expression of *SH3RF3/POSH2* and
211 *FIRRE*, a MYC-regulated long non-coding RNA (lncRNA). Finally, c5 displayed upregulation of
212 immunoglobulin heavy and light chains (*IGHA1, IGHG1, IGHM, IGKC, IGLC1-3*) as well as genes
213 involved in B cell differentiation (*MZB1, PRDM1/BLIMP1, XBP1*). Gene ontology (GO) networks
214 were also generated for top DEGs from one-versus-all-other comparisons to facilitate phenotype
215 annotations (**Figures S12-S16**).
216
217 A map of B cell phenotypes and fate trajectories in early EBV infection
218 Graph-based pseudotime (Qiu et al., 2017) approximated EBV-induced state transitions when
219 anchored from resting cells (**Figure 3A**). Pseudotime scoring was used to track state dynamics
220 of the top 25 marker genes for each phenotype and four example expression trajectories are
221 highlighted (**Figure 3B**). Collectively, flow cytometry for the B cell marker CD19 and CD23
222 (upregulated in EBV infection) at each timepoint (**Figure S17**), cluster-specific DEGs, network
223 ontologies, and pseudotime led us to propose a multi-phenotype model for heterogeneous cell
224 fate trajectories (**Figures 3C, S18-S19**) that manifest in early EBV infection *in vitro*. In this model,
225 naïve (c8) and memory (c3) B cells infected with EBV either undergo antiviral response-mediated
226 arrest (c4) or EBV-driven hyperproliferation (c6) within several days of infection.
227 Hyperproliferating cells can subsequently enter one of several activated states (c0, c1, c2) or
228 undergo growth arrest (c7). Further, differentiated B cells (c5) can develop following activation in
229 a manner analogous to effector cell exit from the germinal center reaction.
230 Among activated phenotypes, c2 matched classical EBV-mediated activation of NF-κB
231 pathway genes, apoptotic regulators, and other known biomarkers (Cahir-McFarland et al., 2004;
232 Messinger et al., 2019). Cells in c1 were consistent with a related activation intermediate that
233 originated from EBV$^+$ naïve cells. Despite the relatively short timecourse, c2 and c5 began to
234 reflect the continuum of activation and differentiation phenotypes we previously characterized in
235 LCLs (SoRelle et al., 2021), which are considered to be immortalized at 21-28 days post-infection
236 (Nilsson et al., 1971). We confirmed these similarities by merging Day 8 and the LCL GM12878,
237 for which scRNA-seq data was previously reported and analyzed (Osorio et al., 2019; Osorio et
238 al., 2020; SoRelle et al., 2021) (**Figure S20**). Conceivably, EBV$^+$ cells could also transition to a
239 plasmablast phenotype (c5) from memory cells (c3) through hyperproliferation (c6) via division-
240 linked differentiation (Hodgkin et al., 1996), effectively bypassing intermediate states.

241 Conversely, cells in c7 highlighted diverse origins of EBV$^+$ cell growth arrest, apoptosis, and
242 senescence, which each provide host defenses against oncogenic malignancies (Bartkova et al.,
243 2006; Nikitin et al., 2010). In addition to highly variable mitochondrial expression and the lowest
244 transcript levels of any state, this phenotype was defined by broad upregulation of genes involved
245 in ribosome biogenesis-mediated senescence (*RPS14*, *RPL29*, *RPS11*, *RPL5*) (Lessard et al.,
246 2018; Nishimura et al., 2015) and stress-associated sequestration of actin monomers that favor
247 G-actin formation (*TMSB4X*, *TMSB10*, *PFN1*) (Kwak et al., 2004). A subset of cells within c7 also
248 contained elevated levels of cell cycle markers (*MKI67*, *TOP2A*, *CCNB1*, *CENPF*) carried over
249 from pre-arrest hyperproliferation (**Figure S21**).
250
251 Evidence for EBV induction of an activated precursor to early memory B cells (AP-eMBC)
252 We next sought to compare early infected phenotypes from our multistate model with cells
253 isolated from secondary lymphoid organs. We acquired single-cell RNA-seq data from human
254 tonsil tissue and identified germinal center (GC) cell subsets (**Figure 4A**), which we analyzed
255 alongside early infection phenotypes of interest. EBV$^+$ NF-κB activated cells (c2) clearly mimicked
256 GC light zone (LZ) B cells; *MKI67*^{hi} cells (c6) matched actively cycling cells (including GC dark
257 zone (DZ) B cells); and EBV$^+$ differentiated cells (c5) matched plasmablasts and plasma cells (PB
258 / PC). Cells in c0 were most like pre-GC naïve and memory B cell (MBC) subsets (**Figure 4B-D**).
259 Further, numerous c0 markers were consistent with both pre-GC activated B cells (*SELL, BANK1,*
260 *CD69, GPR183 (EBI2)*) and memory B cell phenotypes (*SELL, BANK1, GPR183, PLAC8*)
261 recently identified from scRNA-seq of tonsils in response to antigen challenge (King et al., 2021)
262 (**Figure 4C-D**). Cells in c0 further exhibited upregulation of genes with essential roles in B cell
263 activation (*TNFRSF13B/TACI*) (Wu et al., 2000) and germinal center formation (*POU2AF1/OCA-*
264 *B*) (Kim et al., 1996; Luo and Roeder, 1995; Schubart et al., 1996) (**Figure S18**). Moreover, c0
265 displayed elevated *CCR6*, a marker of an activated precursor (AP) state that can generate early
266 memory B cells (eMBCs) (Glaros et al., 2021; Suan et al., 2017) (**Figure 5A**).
267 We subsequently validated the generation of CCR6$^+$ AP-eMBC B cells in response to EBV
268 infection through time-resolved FACS (**Figures 5B-D, S22-S24**). Resting B cells were CCR6^{lo}
269 and remained so until between 2 and 5 days after infection. Further, we observed that the most
270 proliferative cell fraction at day 8 was CCR6^{lo} and a moderately proliferative cell population was
271 CCR6^{hi}. While the most proliferative cells were CCR6^{lo}/CD23^{lo}, the proliferative CCR6^{hi} cells
272 displayed variable CD23 levels (**Figure 5B**). Consistent with our scRNA time course,
273 CCR6^{hi}/CD23^{hi} and CCR6^{hi}/CD23^{lo} populations respectively corresponded to c1/c2 and c0 and
274 emerged within 5 days (**Figure 5C**). Based on CD27 and IgD status, these populations

275 predominantly originated from naïve or non-switched memory versus switched memory cells,
276 respectively; notably, cells from these different resting phenotypes were present in each
277 population gated by CCR6 and CD23 status (**Figure 5D, S24C**). Rapidly proliferative
278 $CCR6^{lo}/CD23^{lo}/CD27^{hi}/IgD^{lo}$ cells were consistent with infected memory B cells transitioning to
279 plasmablasts (*c3→c6→c5* model trajectory; ~72% of $CCR6^{lo}/CD23^{lo}$ cells). Marginally less
280 proliferative $CCR6^{hi}/CD23^{lo}/CD27^{hi}/IgD^{lo}$ cells were consistent with stimulated AP-eMBCs
281 (*c3→c6→c0* model trajectory; ~74% of $CCR6^{hi}/CD23^{lo}$ cells). We also observed an IgD^{hi} naïve
282 population that matched the pre-GC AP-eMBC phenotype (*c8→c6→c0*; (~25% of $CCR6^{hi}/CD23^{lo}$
283 cells). Finally, an even less proliferative $CCR6^{hi}/CD23^{hi}/IgD^{hi}$ population matched activated naïve
284 (or non-switched memory) cells destined for GC BC (*c8→c6/c0→c1/c2*; ~80% of $CCR6^{hi}/CD23^{hi}$
285 cells) and the minor subset (~17%) of $CCR6^{hi}/CD23^{hi}$ cells that was IgD^{lo} was consistent with
286 MBCs induced by EBV to undergo a pseudo-GC reaction (*c3→c6/c0→c2*) (**Figure 5C-E**).
287 Intriguingly, a subset of $CCR6^{+}$ cells displaying the AP phenotype apparently persists long after
288 the early stages of infection based on scRNA-seq data from LCLs (**Figure S25**). Thus, c0 in our
289 model matches a virus-induced common progenitor state from which PBs, GC BCs, and early
290 MBCs have been shown to originate in response to antigen stimulation (Taylor et al., 2015). Our
291 results further indicate that both naïve and memory B cells can achieve this multipotent state at
292 different frequencies upon *in vitro* infection and that the AP-eMBC phenotype is perpetuated in
293 EBV-immortalized B cells.
294
295 Linked expression and accessibility illuminate regulatory mechanisms in phenotype trajectories
296 We next investigated potential regulatory mechanisms underlying DEGs observed across
297 phenotypes. Expression data were jointly analyzed with cell-matched measurements from single-
298 cell Assay for Transposase-Accessible Chromatin sequencing (scATAC-seq) and annotated by
299 state (**Figures 6A-B**, **S26A-B**). While total and unique transcripts per cell increased through early
300 infection, global chromatin accessibility decreased substantially upon infection. Resting (c3, c8)
301 and hyperproliferative (c6) cells had the highest overall accessibility. There were significantly
302 more peaks in the NF-κB activation state (c2) relative to other activation intermediates (c0 and
303 c1; two-tailed t-test, $p < 2.2\times10^{-16}$ and 1.4×10^{-14}, respectively) and differentiated cells (c5; $p <$
304 2.2×10^{-16}). Similar accessibility reduction occurred in both donors in the first five days, with
305 increased accessibility recovered between Day 5 and Day 8 (including to higher than resting
306 levels in one donor) (**Figures 6B, S27**). This indicated that EBV-induced heterochromatinization
307 is likely transient in successfully infected cells (i.e., those that evade innate- and damage-
308 mediated arrest).

309 We found 954 linked feature genes derived from the top 100 marker genes for each cluster
310 using the multimodal integration capabilities of the Signac package in R (Stuart et al., 2021). Of
311 these 954 genes, 177 were significant DEGs with linkages to 476 differentially accessible peaks
312 (DAPs). This translated to 18.6% of tested genes with potential DAP-linked regulation (**Figures**
313 **6C, S26C-D**). We identified genes linked (in *cis*) to DAPs to explore phenotype-associated gene
314 regulatory relationships (**Figure S28A**). Joint analysis of DAP-linked DEGs yielded four regulatory
315 patterns: higher accessibility with higher expression (+/+); lower accessibility with higher
316 expression (-/+); lower accessibility with lower expression (-/-); and higher accessibility with lower
317 expression (+/-). (**Figure S28B**). The +/+ and -/- patterns were characteristic of positive regulatory
318 sites. The less frequently observed -/+ and +/- patterns were consistent with closure or opening
319 of ATAC sites with negative regulatory functions, respectively. Analysis of genes of interest
320 including *CCR7*, *CXCR4*, *RUNX3*, *BACH2*, *JCHAIN*, and *PRDM1* provided examples of each
321 regulatory pattern and their variation among states (**Figures S28C, S29-S30**).
322 We developed joint scRNA + scATAC profiles for major infection fate trajectories in our model
323 (**Figure 6C-F**). The path from resting cells to EBV$^+$ arrested/senescent cells (*c38→c47*) was
324 characterized by global reductions in accessibility and expression. 34.3% of all DEGs between
325 resting cells and these non-proliferative EBV$^+$ fates were linked to DAPs that become inaccessible
326 after infection (*c38!c47* peaks). Top DAP-linked DEGs in innate arrested cells (c4) corresponded
327 to upregulation of interferon-responsive genes and downregulation of mitochondrial genes, while
328 stress- and damage-induced senescent cells (c7) were distinguished by their high expression of
329 interferon-stimulated genes (e.g., *ISG15*) and ribosomal transcripts (e.g., *RPS14*). Cells in each
330 of these clusters also displayed reduced expression of the proto-oncogenic tyrosine kinase gene
331 *LYN* linked to closure of multiple regulatory sites following infection (**Figure 6D**).
332 The reduction in ATAC peaks within c7 was consistent with the formation of senescence-
333 associated heterochromatin foci (SAHF) (Courtois-Cox et al., 2008; Di Micco et al., 2011; Lenain
334 et al., 2017) (**Figure S31**). Because senescence can arise from diverse mechanisms such as
335 innate immune sensing or growth-induced DNA damage, we used higher resolution clustering to
336 reveal c7 subsets (7a and 7b). These subsets displayed DEGs involved in the cell cycle and
337 antiviral sensing (**Figure S32**). Different *HMGB2* levels between 7a and 7b were notable, as this
338 gene's product mediates diverse roles in sensing (Yanai et al., 2009), double-stranded break
339 repair (Krynetskaia et al., 2009), and p53 downregulation (Stros et al., 2002). Relative to resting
340 B cells, *HMGB2* expression was strongly elevated in 7b (as in the hyperproliferative state, c6) but
341 only mildly so in 7a (similar to c4, which could precede senescence (Glück et al., 2017)). Similarly,
342 cell cycle markers were lower in 7a than 7b. Thus, 7a was consistent with EBV$^+$ cells that arrest

343 almost immediately via innate sensing and become senescent, whereas 7b matched a trajectory
344 in which EBV$^+$ hyperproliferative cells become senescent following replicative stress response
345 induction. Notably, both 7a and 7b exhibited elevated levels of ribosomal subunit mRNAs (**Figure**
346 **S32C-D**).
347 NF-κB activated EBV$^+$ cells (c2) exhibited loss of accessibility at 1,142 sites present within
348 resting cells in both donors (8.7% of all resting cell peaks). This reduction paralleled upregulated
349 expression of the polycomb group repressor *EZH2* and a polycomb-interacting methyltransferase,
350 *DNMT1* (**Figure S33**). However, EBV-activated cells possessed 668 peaks absent in resting cells
351 (*c2!c38*) that were linked to 595 unique genes. 154 of these (25.9%) were DEGs between resting
352 and activated cells (**Figure 6E**). These 154 *c2!c38* DAP-linked DEGs included 109 upregulated
353 and 45 downregulated genes from *c38→c2*. Upregulated genes included regulators of apoptosis
354 and tumor suppression (*BCL2A1, TNFRSF8, PDCD1LG2, ST7, IQGAP2, TOPBP1, CD86*);
355 proliferation (*CDCA7, MKI67*); B cell signaling (*NFKBIA, MAPK6, TNIP1, TRAF3*); inflammation
356 (*SLC7A11, RXRA, ZC3H12C*); oxidative stress (*SLC15A4, TXN*); and epigenetic remodeling
357 (*AHRR, NCOR2*). The 45 downregulated *c2!c38* DAP-linked DEGs included *CCR7*,
358 acetyltransferases (*EPC1, KAT6B*), apoptotic and stress response regulators (*STK39,*
359 *STK17A/DRAK1, VOPP1, ZDHHC14*), negative regulators of B cell signaling (*CBLB*), and the
360 tumor suppressor *ARRDC3* (**Figure S34**).
361 In a third example, we explored DAP-linked DEGs between EBV-induced activated (c2) and
362 differentiated (c5) phenotypes. Activated cells exhibited 999 called peaks that were absent in
363 differentiated cells (*c2!c5* DAPs) while only 13 new peaks emerged in differentiated cells (*c5!c2*
364 DAPs) in both donors. This corresponded to a 15% net reduction in accessible peaks in the
365 *c2→c5* transition. Notably, *c2!c5* DAPs found in both donors were linked to 13.4% of all DEGs
366 identified between these states in the scRNA assay. Key *c2→c5* DAP-linked DEG dynamics
367 included downregulation of NF-κB family genes and upregulation of plasmablast-specific
368 transcriptional regulators, translation factors, and protein export machinery (i.e., facilitators of Ig
369 synthesis, secretion, and protein folding chaperones) (**Figures 6F, S10**).
370 By mapping multiome reads to a concatenated reference (human + EBV), we were able to
371 detect increased accessibility within the EBV genome over time after infection. We detected 21
372 unique viral ATAC peaks (20 of 21 common to both donors) including at TSSs for essential viral
373 genes such as the *EBNAs* and *LMP1* (**Figure S35**). Quantification of episome peak-containing
374 cells by phenotype revealed that EBV$^+$ activated and hyperproliferative cells had the greatest
375 number of accessible loci relative to other post-infection phenotypes. These sites included the C
376 promoter (Cp) for *EBNA1*, *EBNA2*, and *EBNA3A-C*; the *LMP1* TSS; the TSS for *BMRF1*, a DNA

377 polymerase accessory protein; and the *BHLF1* locus, which was recently recognized as a
378 facilitator of latency and B cell immortalization (Yetming et al., 2020) (**Figure S36A-B**). Innate
379 arrested cells (c4) exhibited the lowest frequency of cells with accessible episomal loci, followed
380 by growth-arrested cells (c7) (**Figure S36C**).
381
382 Post-infection cell fates exhibit differential enrichment of TF motifs
383 To further investigate regulatory differences by phenotype, we assayed TF binding motif
384 enrichment by state (**Figure S37**). We identified variable motif enrichment linked to resting cell
385 phenotypes and among non-arrested post-infection states (**Figure S37A**). Variation in accessible
386 motifs broadly aligned with phenotypic gene expression with respect to antiviral response
387 induction, promotion of cell growth, and oncogenesis (**Figure S37B-C**). Activated B cell (c2)
388 ATAC peaks were enriched in binding sites for proto-oncogenic TFs including members of the
389 REL (cREL, RelA, RelB), AP-1 (FOS, FOSB, JUNB, JUND), and EGR (EGR1-4) families.
390 Enhanced accessibility at NF-κB family binding sites within activated cells was noteworthy, given
391 the observed concurrent upregulation of NF-κB pathway gene expression. Similar phenotypic
392 consistency was observed within differentiated cells (c5), which were enriched in accessible
393 motifs for IRF4, IRF8, and XBP1. Globally, both resting B cell phenotypes and the innate sensing
394 arrest state shared the greatest motif correlation with each other ($R>0.75$) and the lowest
395 correlation with EBV-activated and hyperproliferating cells ($0.55<R<0.7$) (**Figure S38**).
396
397 An informatics pipeline to infer phenotype-resolved TF signatures and gene regulatory elements
398 The prevalence of DAPs linked to DEGs known to be modulated in *trans* by EBV gene products
399 led us to interrogate phenotype-specific TF signatures genome-wide. To do so, we employed a
400 bioinformatic workflow to obtain ChIP-seq referenced inferences of single-cell phenotypes from
401 scATAC data, which we termed "crisp-ATAC" (**Figures S39, S40**). We expected that ensemble-
402 averaged ChIP data from an appropriate reference cell type would contain TF binding (and
403 epigenetic) data from a superposition of cell phenotypes at high coverage, thus maximizing
404 chances to identify overlaps with comparatively sparse scATAC cluster data. We further reasoned
405 that phenotypic variation in TF binding site accessibility would have biological consequences.
406 We sought to predict viral EBNA and LMP1-mediated NF-κB accessible sites at promoters,
407 enhancers, and actively transcribed genes in each state of our model. To do so, we applied crisp-
408 ATAC recipes to intersect peaks from each scATAC phenotype with ChIP-seq peaks for viral
409 EBNAs, NF-κB/Rel TFs, H3K4me1, H3K4me3, H3K27ac, H3K36me3, and RNA Pol II (Jiang and
410 Mortazavi, 2018) (**Figure S39C**). Hyperproliferative cells (c6), EBV-activated cells (c2), and

411 resting memory B cells (c3) exhibited up to 3-fold more enhancers and promoters at known
412 EBNA2 binding sites relative to naïve B cells (c8), other activation intermediates (c0, c1),
413 plasmablasts (c5), arrested states (c4, c7), and non-B cells (c9, c10). Similar patterns were found
414 for EBNA3C and EBNALP sites (**Figure S39C**, left column) as well as Rel family TF binding sites
415 (cRel, RelA, and RelB). Enhancers, promoters, and actively transcribed genes were consistently
416 enriched in c2, c3, and c6 and depleted in c4 and c7, with intermediate levels present in c0, c1,
417 c5, and c8. By accounting for peaks conserved in both biological replicates, we demonstrated
418 low-noise measurements of DAPs for use with crisp-ATAC and characterized DAP frequencies
419 and interval length distributions across all pairwise phenotype comparisons (**Figure S40**).
420
421 crisp-ATAC finds TF-linked expression signatures that vary across distinct EBV$^+$ cell fates
422 We applied crisp-ATAC to capture regulatory variation among infection phenotypes with
423 respect to key viral transcriptional co-activators. We first compared the innate sensing arrest (c4)
424 and NF-κB (c2) states (**Figure 7A**), as these represent starkly different post-infection fates. Peak
425 data were extracted and gated to obtain *c2!c4* DAPs present in both donors (n=1,873), which
426 yielded linked gene predictions (n=1,514) (**Figure 7B-C**). The *c2!c4* linked gene ontology network
427 was enriched for innate defense (inflammation, antimicrobial processes) and EBV-induced
428 responses (lymphocyte activation, regulation of apoptosis) (**Figure 7D**). Taking a macroscopic
429 view, we found that predicted *c2!c4* linked genes included 42.5% (71 of 167) of known EBV super-
430 enhancer (EBVSE) site-linked genes (Zhou et al., 2015). Consequently, 41-55% of EBNA-
431 associated *c2!c4* DAPs also overlapped a peak for the super-enhancer-associated host TF RelA.
432 Of the 71 EBVSE-linked genes identified for *c2!c4*, 19 (27%) were linked to EBNALP ChIP-seq
433 peaks; 22 (31%) to EBNA2 ChIP-seq peaks; and 15 (21%) to EBNA3C ChIP-seq peaks. EBVSE-
434 linked genes were enriched in EBNA-associated DAP-linked DEGs relative to size-matched
435 random samples of genes in the captured transcriptome (**Figure 7E**).
436 We analyzed specific genes of interest based on 1) EBVSE membership, 2) GO process
437 involvement, and/or 3) empirically demonstrated importance to EBV infection. The NF-κB
438 activated gene and signal transducer *TRAF1*, whose gene product interacts with viral LMP1, was
439 identified through all three of these routes (Devergne et al., 1996; Eliopoulos et al., 2003; Fries et
440 al., 1999; Greenfeld et al., 2015; Sandberg et al., 1997). We found *c2!c4* DAPs associated with
441 one or more EBNA at -3kb, +2kb, and +37kb relative to the *TRAF1* TSS, each with significant
442 positive correlation to *TRAF1* expression ($p<0.05$ for correlation z-score). Notably, these EBNA-
443 associated regulatory loci exhibited reduced accessibility in stress arrest (c7), activation
444 intermediate (c1) and differentiated (c5) states compared with c2 (**Figure 7F**).

445 We used a grouped crisp-ATAC comparison (*c256!c38*) to study changes associated with the
446 trajectory for successful EBV-induced B cell immortalization (**Figure S41**). We analyzed viral co-
447 activator-associated DAPs between proliferative (c6) and LCL-like phenotypes (c2, c5) versus
448 resting B cells (c3, c8). Despite the net reduction in accessibility after infection, we identified 245
449 unique genes linked to 1,824 peaks present in all tested EBV$^+$ states (c256) but absent in both
450 resting phenotypes (c38) (**Figure S41A-C**). 166 of the 245 genes were linked to a binding site for
451 at least one EBNA, and 18 of these genes overlapped with EBVSE targets (7.3% of predicted
452 genes, 10.8% of known EBVSE genes). Only 31 GO process terms were shared across the top
453 100 terms for each tested EBNA, accounting for 15% of unique terms (**Figure S41D**). We selected
454 the *c256!c38*∩EBNA targets *TNFRSF8* (CD30), *CD274* (PD-L1), and *PDCDL1G2* (PD-L2) based
455 on their therapeutic relevance to EBV-associated lymphomas. For each gene, we confirmed the
456 presence of EBNA-associated DAP-linked DEGs by phenotype. These included three
457 $EBNA2_{c256!c38}$ sites near the *TNFRSF8* TSS (-17kb, -12kb, and +16kb) and a shared multi-EBNA
458 site -17kb from the *PDCDL1G2* TSS and +43kb from the end of the *CD274* gene (**Figure S41E**).
459 These loci were enriched for Rel sites and activating histone marks in LCL reference data.
460 In a final example, we evaluated activated versus differentiated EBV$^+$ phenotypes (*c2!c5*) sites
461 with known viral transcriptional co-activator binding in both donors to explore regulatory
462 relationships that distinguished the phenotypes present in LCLs (SoRelle et al., 2021) (**Figure**
463 **S42A**). 519 of 999 identified *c2!c5* peaks intersected with at least one EBNA binding site, from
464 which 247 unique genes were predicted. 29 of 110 *c2!c5*∩EBNA2 site-linked genes (26.3%) were
465 *c2!c5* DEGs in the scRNA assay, as were 34 of 115 *c2!c5*∩EBNALP site-linked genes (29.6%)
466 and 20 of 125 *c2!c5*∩EBNA3C site-linked genes (16.0%). 20 site-linked genes were identified
467 from all three viral co-activator recipes (**Figure S42B**), including the EBVSE-linked G protein
468 coupled receptor *GPR137B*, a lysosomal transmembrane receptor that regulates mTORC1 activity
469 and autophagy (Gan et al., 2019; Gao et al., 2012). *GPR137B* was also identified as a *c2!c38*
470 DAP-linked DEG, indicating inaccessible regulatory loci within resting cells as well. We identified
471 two regulatory DAPs with significant positive correlation to gene expression at +14kb and +18kb
472 relative to the *GPR137B* TSS. One of these sites (+14kb) coincided with EBNALP and EBNA3C
473 binding sites but did not exhibit a *c2!c5* DAP. However, the second site (+18kb) overlapped with
474 all three intersected EBNAs and was a *c2!c5* DAP. Both sites also intersected with Rel family TFs
475 (cRel, RelA, and RelB) (**Figure S42C**). Other genes involved in lysosome-mediated processes
476 including autophagy and antigen presentation regulation were also identified from the *c2!c5*
477 comparison (*TFEB*, *LAMP3*), albeit with modestly elevated but significant differential expression
478 (Martina et al., 2012; Nagelkerke et al., 2014; Settembre et al., 2011). We also found numerous

479 other genes involved in immune activation signaling, apoptosis, and transcriptional regulation that
480 displayed EBNA-associated *c2!c5* DAP-linked DEGs (**Figure S42D-E**). Collectively, these
481 vignettes of post-infection cell trajectories highlight the genome-wide dynamics of diverse EBV-
482 induced B cell responses captured within the single-cell multiomics dataset.

483 **Discussion**

484

485 Our data reveal heterogeneity in coordinated gene expression and chromatin accessibility
486 dynamics within individual cells during the critical early stages of a viral infection. By capturing
487 the initial phases of EBV infection with high resolution -omics, we discern the gene regulatory
488 environments associated with diverse infected cell fates and their respective developmental
489 trajectories. These include genome-wide expression and chromatin signatures associated with
490 effective host antiviral response, virus-triggered oncogene-induced senescence, and the path to
491 sustained EBV latency and host cell immortalization, which is accessed via simulated B cell
492 activation. Moreover, we develop a bioinformatic workflow to characterize post-infection
493 outcomes through gene- and peak-level investigations at loci matching specific epigenetic
494 patterns as well as host and viral transcription factor binding profiles. The combined high
495 resolution multiomics data and integrative analytical framework reported herein yield a vividly
496 detailed representation of the genome-wide interplay of host and virus.

497 We captured diverse post-infection B cell fate trajectories that, due to their asynchronous
498 parallel emergence, cannot be resolved by ensemble sequencing methods. Remarkably, we find
499 that large-scale euchromatin-to-heterochromatin transitions (20-40% reductions in genome-wide
500 accessibility) can occur in post-infection trajectories and fates, including several that display
501 increases in transcript levels. The scATAC data implicate Simpson's paradox with respect to
502 heterogeneous chromatin accessibility dynamics following infection, since the total number of
503 unique peaks increases in early infected populations while the peaks per cell can, in fact,
504 decrease (Simpson, 1951; Trapnell, 2015) (**Table S1**).

505 Global epigenetic silencing via SAHF formation (Di Micco et al., 2011) is most prominent both
506 in cells that rapidly undergo innate sensing-mediated arrest and in cells that evade this response
507 but arrest due to DNA damage response activation during virus-induced hyperproliferation. Our
508 data also indicate a possible role for ribosome biogenesis in the transition from virus-induced
509 arrest to senescence, likely through a p53-MDM2 axis (Deisenroth and Zhang, 2010), which
510 critically regulates EBV transformation (Forte and Luftig, 2009). Intriguingly, the trajectory of
511 successful infection is distinguished by increased accessibility at key sites against the larger
512 backdrop of heterochromatin formation. A substantial number of these sites (many associated
513 with viral TFs) have predicted *cis* linkages to genes enriched for regulation of apoptosis, tumor
514 suppression, inflammation, and chromatin remodeling, all of which are determinants of successful
515 EBV infection. Moreover, viral episome ATAC profile heterogeneity across infected phenotypes

516 indicates that latency establishment depends on retained accessibility to viral genes within the
517 repressive host milieu.
518 The identification of an EBV-infected analog of an AP-eMBC phenotype is consistent with
519 results from *in vitro* and *in vivo* antigen stimulation experiments (Suan et al., 2017; Taylor et al.,
520 2015), as well as previous work from our lab that identified CCR6 as an EBV Latency IIb program
521 biomarker that becomes downregulated in the transition to Latency III in LCLs (Messinger et al.,
522 2019).The development of this state *in vitro* implies that EBV may gain access to the memory
523 pool *in vivo* via progenitors that have limited involvement in the GC reaction. In the context of
524 normal antigen stimulation, this subset of eMBCs undergoes early exit from the cell cycle and GC
525 reaction as a consequence of restricted access to or engagement with cognate antigen (Glaros
526 et al., 2021). It is thus conceivable that EBV-infected B cells may differentially develop into GC
527 BCs, PBs, or eMBCs from the AP state in a manner dependent on the extent to which the LMPs,
528 EBNAs and other viral gene products induce mimicry of BCR activation and signaling. This model
529 accommodates a surprising possibility – that EBV may access long-term persistence and survival
530 not only within post-GC high-affinity MBCs but also via GC-independent eMBCs that avoid
531 extensive proliferation.
532 The activated B cell and plasmablast phenotypes that developed within 5 days are generally
533 consistent with our findings in LCLs (SoRelle et al., 2021) and resemble *in vivo* tonsil cell subset
534 transcriptomes. Furthermore, differentiated plasmablasts exhibited fewer accessible sites in both
535 host chromatin and viral episomes relative to activated cells. Notably, prior studies found that only
536 50% of EBV-infected cells that secrete immunoglobulin go on to become immortalized (Tosato et
537 al., 1985). Additional work within LCLs demonstrated that EBV^+ cells with upregulated Ig
538 production exhibited reduced DNA synthesis and EBNA downregulation (Wendel-Hansen et al.,
539 1987). Collectively, these findings support a model for continuous EBV-driven B cell differentiation
540 *in vitro*, wherein plasmablasts are continually generated through activation-induced maturation
541 yet selected against by their reduced proliferation. While this disadvantage limits viral replication
542 via cell division, the reduced MYC levels, increased XBP1, and endoplasmic reticulum stress in
543 these cells may support EBV lytic reactivation (Guo et al., 2020; Laichalk and Thorley-Lawson,
544 2005; Sun and Thorley-Lawson, 2007; Taylor et al., 2011).
545 Depending on the phenotype comparison, roughly 10-35% of all DEGs were correlated with
546 DAPs. This range was similar to the frequencies of both differential accessibility of expressed TF
547 binding motifs (23%) and DEGs associated with differentially accessible promoters (25%)
548 identified in dexamethasone-treated A549 cells profiled with sci-CAR (Cao et al., 2018). The
549 observation of similar DAP-linked DEG frequencies in response to diverse stimuli (e.g., chemically

550 induced glucocorticoid receptor activation, viral infection) raises intriguing questions regarding the
551 fundamental frequency of genes regulated, at least in part, by accessibility changes. The
552 observed proportion implies that most DEGs may be regulated by higher order chromatin
553 interactions and differential recruitment of transcriptional activators and/or RNA Pol II.
554 The crisp-ATAC method provides a simple, flexible informatic approach to map ChIP-seq
555 profiles to cell phenotypes discovered from scATAC-seq. Thus, it provides a potential route for
556 effectively bootstrapping scATAC resolution to datasets from other -omics modalities. When data
557 from suitable reference cell types and/or states are available, crisp-ATAC can be used to predict
558 phenotype-resolved gene regulation by evaluating simple or complex combinations of factor
559 binding sites, histone modifications, and/or nuclear chromatin compartmentalization across all
560 differentially accessible sites. In our case, cell-matched empirical scRNA-seq data can be cross-
561 referenced with crisp-ATAC outputs for methodological validation and identification of genes of
562 interest for future studies. This method is adaptable to comparisons of phenotype-resolved ATAC
563 profiles in contexts such as other infections, development, and cell responses to stimuli or
564 therapies. We expect this approach to be particularly powerful for exploring TF-associated
565 regulatory changes in time-resolved studies of cellular behaviors.
566
567 Limitations of the study
568 The reported single-cell multiomics data have several limitations. They do no capture aspects
569 of host-virus dynamics acting at other molecular levels. Examples include epigenetic
570 modifications (e.g., DNA methylation status), three-dimensional chromatin architecture changes,
571 modulation of translation and protein abundance, post-translational modifications, protein-protein
572 interactions, and signaling cascades (e.g., phosphorylation status). While we present a method
573 for inferring DAP-linked TF binding and epigenetic modifications based on empirical scATAC data,
574 we do not have direct ChIP-seq measurements at the single-cell level.
575 The reported bioinformatic methodology (crisp-ATAC) also has notable constraints. This
576 approach is limited by the availability of ChIP-seq data from an appropriate cell type and target
577 reference state. Moreover, regulatory site predictions must be empirically tested to validate
578 potential functions in gene expression control. Finally, distance limits imposed for identifying *cis*-
579 regulatory linkages preclude identification of distal gene regulatory elements formed via 3D
580 nuclear conformation.
581

Acknowledgments

The authors wish to thank members of the Duke University Molecular Genomics Core (MGC) and the Duke Center for Genomic and Computational Biology (GCB), especially Emily Hocke, Karen Abramson, Dr. Simon Gregory, and Dr. Nicolas Devos. Special thanks are also due to the anonymous donors, without whose blood donations this work would not have been possible. E.D.S. wishes to acknowledge funding support from the Department of Molecular Genetics and Microbiology Viral Oncology Training Grant (NIH T32 # 5T32CA009111-42). This work was supported by funding from the National Institute of Dental and Craniofacial Research (NIDCR award #R01DE025994-06).

592 **Experimental Methods**

593

594 PBMC isolation and B lymphocyte enrichment

595 Whole blood (50 mL each from two anonymous donors; TX1241/Donor 1 & TX1242/Donor 2)
596 was obtained from the Gulf Coast Regional Blood Center (Houston, TX). Upon receipt, peripheral
597 blood mononuclear cells (PBMCs) from each donor sample were isolated via Ficoll gradient
598 separation (Histopaque®-1077, Sigma #H8889), resuspended at $10x10^6$ cells/mL in RPMI 1640
599 + 15% heat-inactivated fetal bovine serum (FBS, v/v, Corning) (R15 media), and incubated
600 overnight at 37°C and 5% CO_2. The next day, $CD19^+$ B cells were enriched from donor PBMCs
601 via negative isolation (BD iMag Negative Isolation Kit, BD Biosciences #558007) and
602 resuspended at $2x10^6$ cells/mL in R15 supplemented with 2 mM L-glutamine, 100 units/mL
603 penicillin, 100 µg/mL streptomycin, and 0.5 µg/mL cyclosporine A ($R15^+$ media). Roughly $45x10^6$
604 B cells were recovered per donor post-enrichment. Following $CD19^+$ validation (see below: Flow
605 cytometry), enriched B cell aliquots (1-2 mL at $3x10^6$ cells/mL) were viably frozen in 90% FBS +
606 10% DMSO and stored in liquid N_2.

607

608 EBV infection and cell culture

609 The type 1 EBV strain B95-8 was obtained in-house as viral supernatant from the inducible
610 B95-9 Z-HT cell line as reported previously (Johannsen et al., 2004). Immediately after
611 withholding and cryopreserving uninfected enriched B cells for each donor (day 0 samples), the
612 remaining cells in culture were infected with B95-8 via resuspension in viral supernatant (100 µL
613 per $1x10^6$ cells) for 1 h at 37°C and 5% CO_2. Infected B cells from each donor were rinsed with
614 1x PBS, resuspended in $R15^+$ media, and incubated at the conditions described above throughout
615 the course of infection. Aliquots were taken from each infected donor culture at 2-, 5-, and 8-days
616 post-infection and viably frozen as described for uninfected day 0 samples.

617

618 Flow cytometry

619 The extent of B cell enrichment was quantified for each donor using flow cytometry. Following
620 negative isolation, cell samples ($2x10^5$ per donor) were rinsed with FACS buffer (1x PBS + 2%
621 FBS), stained with phycoerythrin (PE)-conjugated mouse anti-human CD19 (BioLegend, clone
622 HIB19; catalog #302208; lot #B273508) in the dark for 30 min at room temperature, then rinsed
623 again prior to analysis. Cell samples at each timepoint were prepared as described and co-stained
624 with PE-anti-CD19 and allophycocyanin-conjugated mouse anti-human CD23 (APC-anti-CD23,
625 BioLegend, clone EBVCS-5; catalog #338513; lot #B273489) to validate successful EBV

626 infection. To validate the AP-eMBC phenotype (c0), enriched resting B cells from two additional
627 donors (TX1253 and TX1254) were labeled with CellTrace Violet (ThermoFisher / Invitrogen, Cat
628 #34571) and stained with one of the following combinations at days 0, 2, 5, and 8: CCR6/Memory
629 panel (FITC-anti-CD27, PE-anti-CCR6, and APC-anti-CD23); Naïve/Memory panel (FITC-anti-
630 IgD, PE-anti-CD19, and APC-anti-CD27); or CCR6/Naïve panel (FITC-anti-IgD, PE-anti-CCR6,
631 and APC-anti-CD23). FITC-anti-CD27, FITC-anti-IgD, and APC-anti-CD27 were purchased from
632 BioLegend (Cat #302806, #348206, and #356410, respectively) and PE-anti-CCR6 was
633 purchased from Invitrogen (Cat #12-1969-42). Compensation matrices were calculated from
634 single-stain controls for FITC and PE and applied to all samples for analysis. All cytometry
635 measurements were acquired with a BD FACSCanto II (BD Biosciences) and analyzed using
636 FlowJo version 10 (Ashland / BD Biosciences).
637
638 Human tonsil sample preparation
639 Tonsillar B cells were isolated from discarded, anonymized tonsillectomies from the Duke
640 Biospecimen Repository and Processing Core (BRPC; Durham, NC). Tonsil tissue samples were
641 manually disaggregated, filtered through a cell strainer, and isolated by layering over a cushion
642 made from Histopaque-1077 (H8889; Sigma-Aldrich). Harvested lymphocytes were washed three
643 times with FACS buffer (5% FBS in PBS) prior to scRNA library preparation.
644
645 Preparation of scRNA and scATAC libraries
646 Cryopreserved samples from each early infection timepoint of interest were simultaneously
647 thawed by donor and purified to > 90% viable cells by Ficoll gradient separation. Viable cells from
648 each timepoint and donor were then prepared as single-cell matched gene expression (scRNA)
649 and chromatin accessibility (scATAC) libraries by the Duke Molecular Genomics Core staff with
650 the 10x Chromium Next GEM Single Cell Multiome ATAC + Gene Expression Kit (10x Genomics,
651 Pleasanton, CA) (Satpathy et al., 2019; Zheng et al., 2017). Briefly, nuclei were isolated from
652 each sample and subjected to transposition at accessible chromatin sites. Next, transposed
653 nuclei, barcoding master mix, and gel beads containing unique barcode sequences were
654 prepared into single-cell GEMs (Gel bead emulsions) using the Chromium Controller and Chip J.
655 Within each GEM, poly-adenylated (poly-A) mRNA transcripts from individual nuclei are captured
656 by barcoded, indexed poly-T primers and reverse transcribed into cDNA. Simultaneously, a
657 separate barcoded sequence containing a spacer and Illumina P5 adaptor sequence is added to
658 transposed regions within the captured nucleus. Barcoded multiomes were then purified, pooled,
659 and pre-amplified by PCR prior to library construction. The scATAC library for each sample is

660 generated by PCR amplification and incorporation of sample index and Illumina P7 adaptor
661 sequences. Separately, pre-amplified gene expression cDNA is further PCR amplified,
662 fragmented, and size selected. The scRNA library for each sample is then constructed using PCR
663 to incorporate the P5 and P7 sequencing adaptors. Two biological replicates of tonsillar
664 lymphocytes were prepared as scRNA libraries using the 10x Genomics Next GEM 3' Gene
665 Expression kit with v3 chemistry (10x Genomics, Pleasanton, CA), and sequenced, processed,
666 aligned, and analyzed as described above for early infection scRNA samples.
667
668 Sequencing, alignment, and count matrix generation
669 The 8 paired-end scATAC libraries (4 timepoints per 2 donors) were pooled onto two lanes of
670 an Illumina S2 flow cell and sequenced at a target depth of 25,000 reads per cell on an Illumina
671 NovaSeq (Illumina, San Diego, CA). The 8 paired-end scRNA libraries were similarly pooled and
672 sequenced at a target depth of 50,000 reads per cell. Tonsil scRNA libraries were likewise pooled
673 and sequenced at 50,000 reads per cell. All sequencing runs were performed by staff at the Duke
674 Center for Genomic and Computational Biology. Raw base calls for each assay were prepared
675 as sample-demultiplexed FASTQ files using *cellranger-arc mkfastq* (Cellranger, 10x Genomics),
676 a wrapper of the Illumina *bcl2fastq* function. Next, sample-matched scRNA and scATAC reads
677 were aligned against genomic references to produce multiome count matrices using *cellranger-*
678 *arc count*. One set of count matrices was generated by mapping reads to a concatenated genomic
679 reference constructed from the human genome (GRCh38) with the ~172 kB type 1 EBV genome
680 (NC_007605) included as an extra chromosome. These outputs were used for downstream RNA-
681 only analyses to capture host and viral gene expression. A second set of count matrices
682 generated by mapping to GRCh38 only was used for downstream joint RNA and ATAC analyses.
683 Compatible reference packages were assembled from the relevant genome (.fa) and annotation
684 (.gtf) files using *cellranger-arc mkref*.
685
686 Data QC and scMultiome analysis
687 All direct analysis of scRNA and scATAC data was conducted in R using Signac (Stuart et al.,
688 2021), an extension of Seurat (Macosko et al., 2015; Satija et al., 2015; Stuart et al., 2019).
689 Following read mapping and counting, linked scRNA and scATAC data were obtained from
690 between 8,934-20,000 cells per sample. After QC filtering by mitochondrial expression (n < 20%),
691 total transcripts (n < 25,000), unique transcripts (n > 1,000), and minimum cells expressing a
692 given feature (n > 3), data from between 8,376-19,310 cells per sample were analyzed (see **Table**
693 **S1**). The mitochondrial gene expression threshold was selected based on the high metabolic

694 activity of early-infected B cells and the high cell viability observed in each sample (> 90%)
695 immediately prior to library preparation to preserve biologically relevant phenotypes (Osorio and
696 Cai, 2021). After QC filtering, a total of 52,271 and 44,920 cells were analyzed across the infection
697 timecourse for donors TX1241 and TX1242, respectively. For RNA-only analysis, gene
698 expression data (host and viral) across all timepoints for a given donor were merged into a single
699 object, log normalized, scored for cell cycle markers, and scaled with cell cycle scoring regressed
700 out. The top 2,000 differentially expressed features over the early infection timecourse data were
701 identified and used for principal component analysis (PCA). The top 30 principal components
702 were further dimensionally reduced via uniform manifold approximation projection (UMAP,
703 (McInnes et al., 2018)), and clustering was performed to identify biologically distinct cell
704 subpopulations. Merged scRNA dataset pseudotime trajectories were calculated using Monocle3
705 (Qiu et al., 2017), and were mapped along with cluster identities to 3D UMAP coordinates for
706 visualization (Qadir, 2019; Qadir et al., 2020). For joint ATAC and RNA analysis, host gene
707 expression and chromatin accessibility were analyzed for each separate timepoint and for a
708 merged object containing Day 0 and Day 8 multiome data. Nucleosome signal and transcription
709 start site (TSS) enrichment were calculated and used for QC filtering (Nucleosome.signal < 2,
710 TSS.enrichment > 1). ATAC peaks were called using macs2 (Liu, 2014) with hg38 annotations.
711 Gene expression data in the joint analysis was processed as described for the RNA-only analysis
712 with the exception of using Signac's SCTransform function instead of log normalization for
713 expression counts. Top differential features in each assay ('peaks' and 'SCT') were determined,
714 and multimodal neighboring and UMAP were performed for integrated data visualization. Cluster
715 identities defined in the RNA-only assay were mapped to this merged joint dataset, which
716 contained cells representing all identified subpopulations. Peaks with significant (anti-) correlation
717 ($p < 0.05$ for z-score of correlation coefficients) to differentially expressed genes were identified
718 using the LinkPeaks function in Signac, which was informed by SHARE-seq (Ma et al., 2020).
719
720 crispATAC workflow and reference data curation
721 ChIP-referenced inference from single-cell phenotype ATAC (crispATAC) was developed to
722 predict subpopulation-resolved gene regulatory features. In a typical workflow, cluster-level
723 chromatin accessibility tracks are cross-referenced against ChIP-Seq (Chromatin
724 Immunoprecipitation Sequencing) profiles for epigenetic marks and TFs of interest measured from
725 a reference cell phenotype (in this study, lymphoblastoid cell lines such as GM12878). In this
726 study, cluster-specific called peaks from the joint scATAC + scRNA dataset were extracted and
727 prepared as simplified genomic range files (3-column .bed file format). Next, the desired ChIP-

728 Seq datasets for the reference phenotype were downloaded and, where applicable, converted to
729 .bedgraph format to be used as input for peak calling with the *macs2* function *bdgpeakcall*. The
730 ChIP (and Hi-C) datasets used for crispATAC in this study (**Table S2**) are all publicly available
731 from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO).
732 Once all ATAC-seq and ChIP-seq peak files were generated, all were used as inputs to a single
733 call of the bedtools (Quinlan and Hall, 2010) function *multiinter*, which output a matrix of all
734 genomic range intersection intervals where at least one input file exhibited a peak. This
735 intersection matrix was imported to R as a data frame and analyzed to identify common and/or
736 differential intervals (matrix rows) among scATAC cluster phenotypes, epigenetic marks, and TFs
737 using Boolean logic gating by dataset (matrix columns). For a given crispATAC recipe (e.g., peaks
738 in scATAC cluster 1 not in scATAC cluster 2 intersected with EBNA2 ChIP peaks = [c1 ∩ !c2] ∩
739 EBNA2), the genome intervals matching the gating criteria were returned and converted to .bed
740 files. Lists of differentially accessible, transcription-factor associated sites generated in this way
741 were subsequently analyzed with the Genomic Regions Enrichment of Annotations Tool (GREAT)
742 (McLean et al., 2010) to identify potential *cis*-regulated genes within 1 megabase of each query
743 site. As a final step, output lists of potential linked genes were intersected with the top marker
744 genes identified from the corresponding cluster-wise comparison in the scRNA assay, thus
745 integrating direct single-cell RNA and ATAC measurements with subpopulation-resolved
746 regulatory inferences from ensemble ChIP profiles. In a similar but separate approach, scATAC
747 and ChIP peaks were intersected with topologically associated domain (TAD) boundaries
748 (prepared using hicExplorer, (Ramírez et al., 2018; Wolff et al., 2018; Wolff et al., 2020)) and
749 nuclear subcompartments from GM12878 Hi-C data to study differentially accessible TF-
750 associated sites in the context of 3D nuclear architecture.
751
752 Visualization of crispATAC outputs, gene ontologies, and networks
753 Data for genes of interest identified from crispATAC recipes were explored using dimensionally
754 reduced (UMAP) expression maps and cluster-level accessibility tracks (Signac, (Stuart et al.,
755 2021)), called peaks aligned with TFs and epigenetic marks (IGV, (Robinson et al., 2011)), and
756 local neighborhoods in Hi-C contact maps (Juicebox, (Durand et al., 2016)). Cluster-resolved
757 gene ontologies were generated and quantified by GREAT (McLean et al., 2010). Top scRNA
758 assay cluster markers and GREAT output gene lists were also visualized as annotated networks
759 using Cytoscape (Shannon et al., 2003).

Main Figures

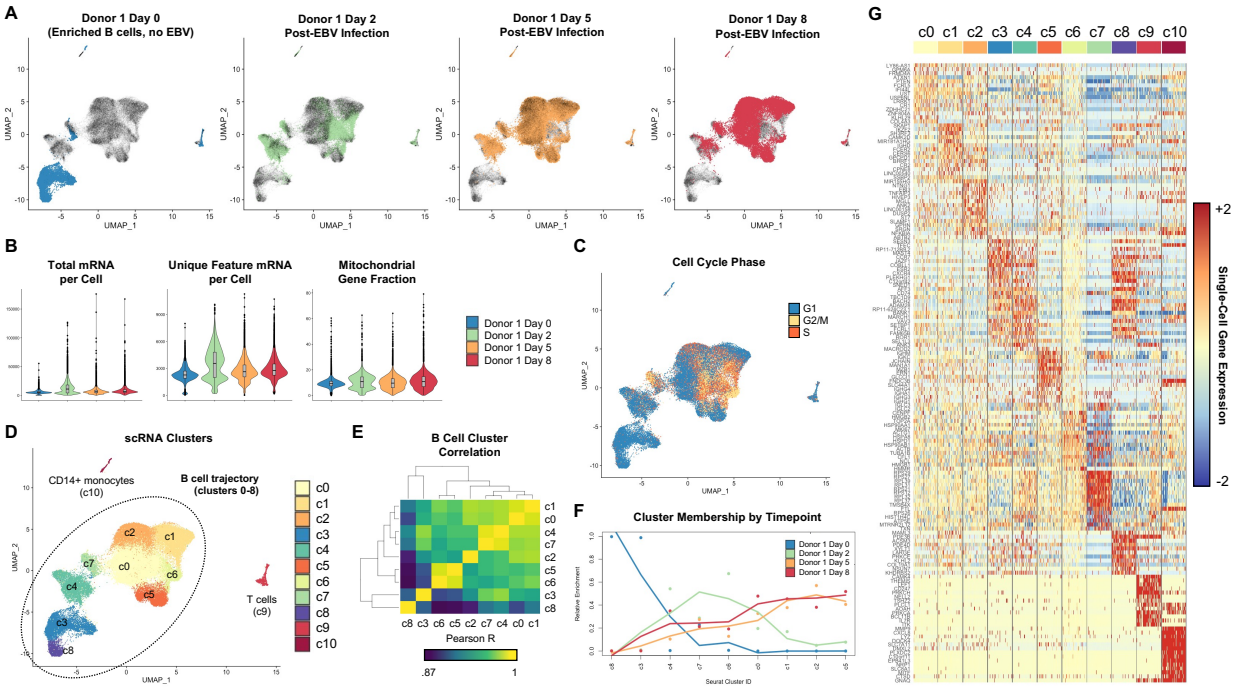

Figure 1. Time-resolved single-cell gene expression during early EBV infection of B lymphocytes

(A) Single-cell gene expression timecourse data from early EBV infection.
(B) General expression trends during early infection. Total mRNA refers to all transcripts captured, while Feature mRNA refers to the number of unique transcripts (per cell).
(C) Cell phase scoring of expression data after cell cycle marker regression.
(D) Unsupervised clustering of early infected cell expression in merged timepoint data.
(E) Pairwise correlation of identified clusters.
(F) Cluster membership by timepoint. Fit lines show coarse changes in phenotype frequency over time.
(G) Single-cell expression of the top 15 gene markers by cluster.
See also Figures S1-S4

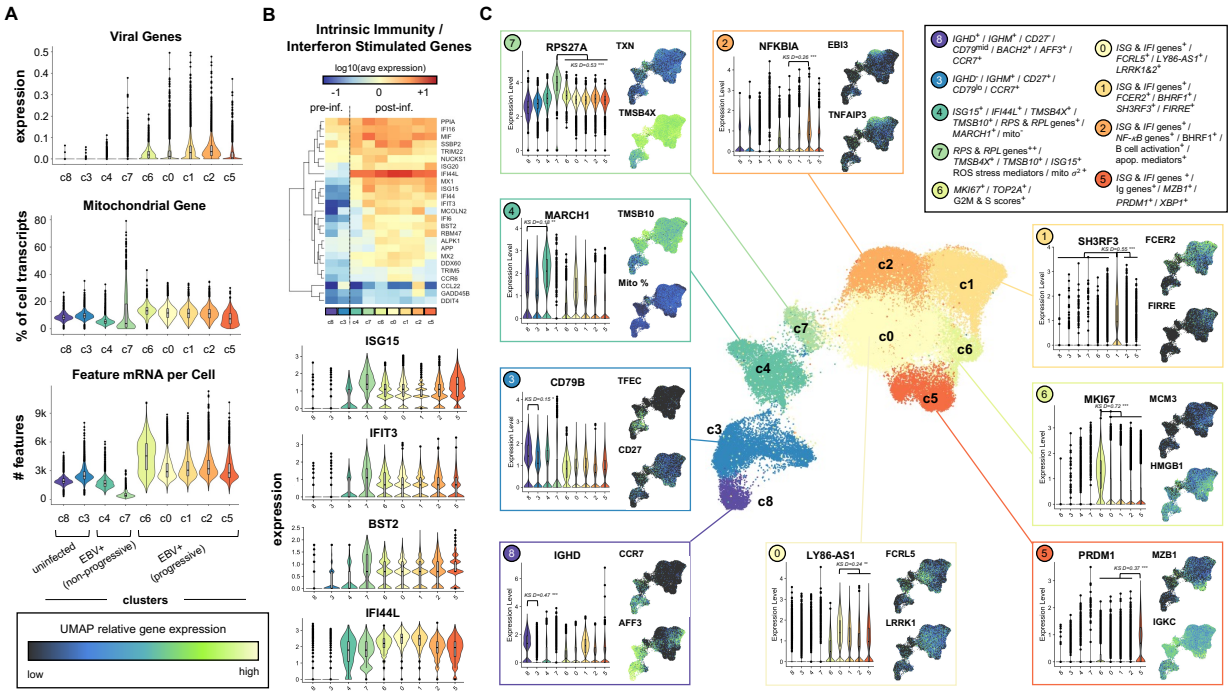

Figure 2. High-resolution dissection of infected B cell phenotypes

(A) Overview of global gene expression trends by phenotype.

(B) Induction of interferon response genes in all EBV$^+$ clusters.

(C) Phenotype-resolved transcriptomic signatures in resting and EBV$^+$ B cells. Select cluster-resolved comparisons of gene expression were evaluated via the Kolmogorov-Smirnov D statistic (KS D) and associated p value (* p < 1e-5; ** p < 1e-10; *** p < 1e-15) from 500 randomly sampled cells per cluster.

See also Figures S5-S11

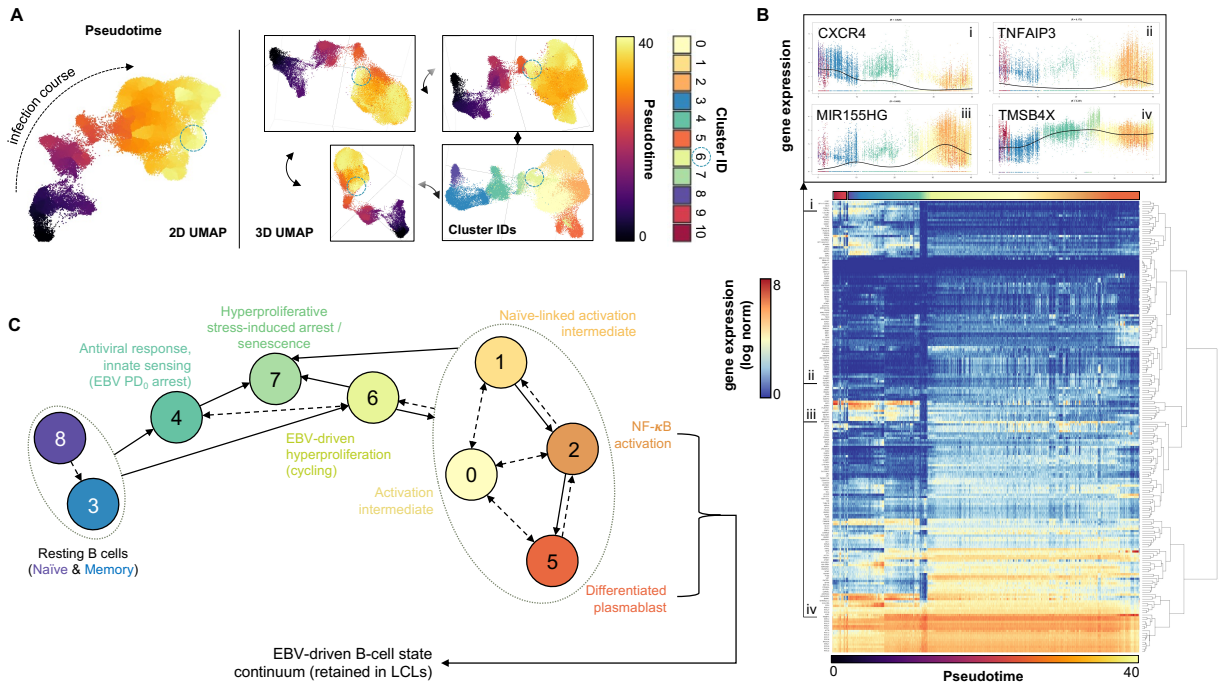

Figure 3. A model of B cell fate trajectories in early EBV infection

(A) Monocle3 pseudotime scoring of merged timecourse expression data relative to resting B cells (day 0). Unlike the 2D UMAP, 3D UMAPs depict closer proximity of c6 (first observed at day 2, blue dashed circle) to resting cells, consistent with the temporal emergence of the c6 phenotype prior to the c0, c1, c2, and c5 phenotypes.

(B) Pseudotime-resolved expression dynamics of top differentially expressed genes (DEGs) across phenotypes. Genes are hierarchically clustered by pseudo-temporal expression pattern similarity. Spline interpolant fits are shown for expression of select genes in pseudotime (insets i-iv). After sorting cells by pseudotime score in ascending order, the average pseudotime score of every 25-cell interval was computed for efficient visualization (i.e., pseudotime for 52,271 cells at 25 cell resolution).

(C) Annotated state model of EBV$^+$ B cell phenotypes and fate trajectories. Empirically observed and putative directed state transitions are depicted in solid and dashed edges, respectively. Edges drawn to groups of phenotypes (dotted ovals) indicate transitions to/from each cluster within the group.

See also Figures S12-S19

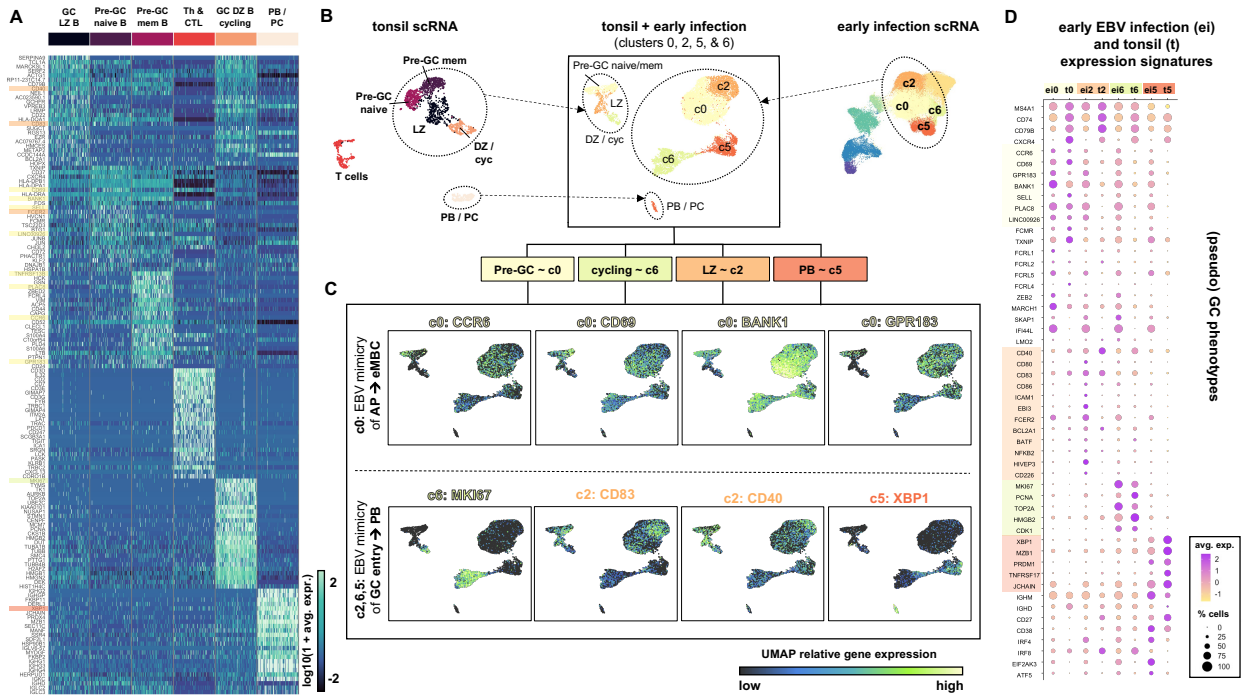

Figure 4. A subset of early infected cells exhibits hallmarks of a multipotent activated precursor to early memory B cells (eMBCs)

(A) Top phenotype markers of healthy human tonsil subsets identified from scRNA-seq.

(B) UMAP merging of tonsil and key early infection cluster scRNA-seq assays. Tonsil clusters are colored to match the closest corresponding cells from early infection.

(C) UMAP Correspondence of key gene expression across tonsillar subsets and early infection phenotypes. Select markers of multipotent progenitors and eMBCs were informed by data from (Suan et al., 2017) and (Glaros et al., 2021).

(D) Dot plot visualization of key genes across early infection (ei) c0, c2, c5, and c6 and their analogs within tonsils (t).

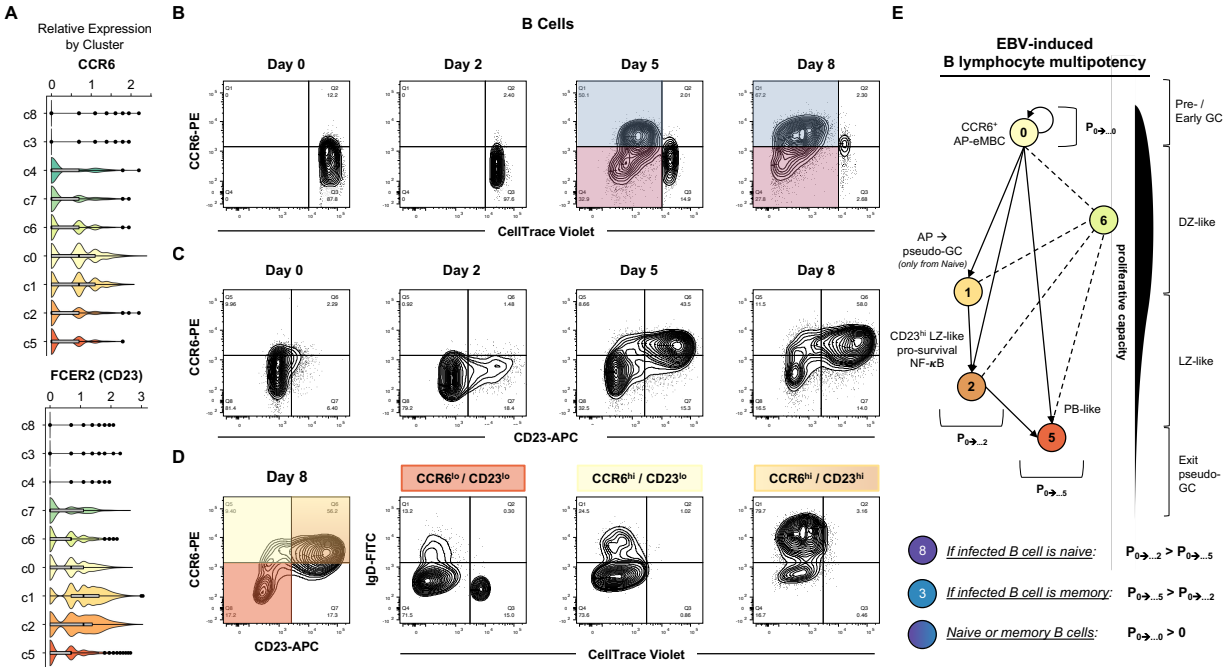

Figure 5. FACS validation of CCR6^{+} AP-eMBCs derived from naïve and memory B cells.
(A) Relative expression of *CCR6* and *FCER2/CD23* by model phenotype determined by scRNA-seq.
(B) CCR6 surface expression on uninfected and EBV-infected B cells by number of cell divisions (CellTrace Violet). CCR6^{hi} cells (blue gate) exhibits reduced proliferation relative to CCR6^{lo} (magenta gate) cells.
(C) CCR6 and CD23 (FCER2) surface expression over the early infection timecourse.
(D) Cell divisions and IgD status of cells at Day 8 gated by CCR6 and CD23 expression. Gated fractions are colored by approximate correspondence to scRNA-seq model phenotypes.
(E) A fate model for EBV-induced AP-eMBC-like cells. $P_{0\rightarrow...n}$ signifies the probability of the eventual transition of a cell from cluster 0 (AP-eMBC analog) to a given cluster n. Relative probability relationships for naïve and memory cells are proposed based on empirical findings from FACS.
See also Figures S22-25

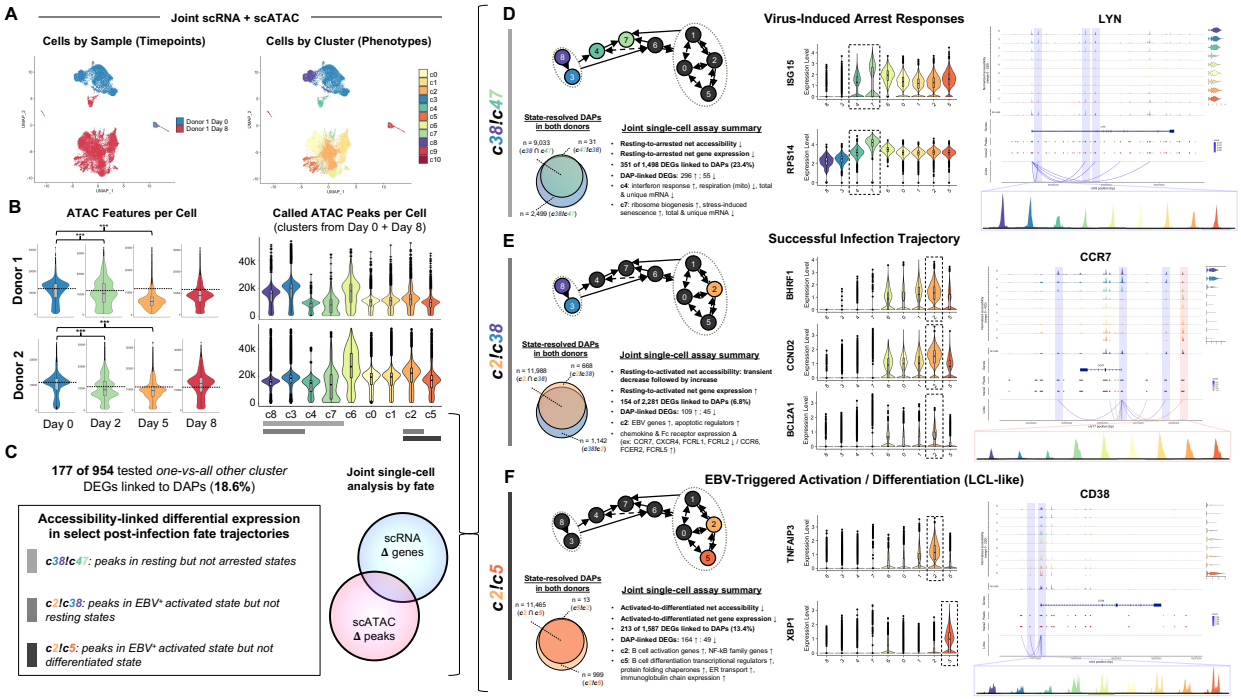

Figure 6. Cell-matched expression and chromatin accessibility cell fate trajectories

(A) UMAP visualization of scRNA + scATAC data generated using weighted nearest neighbors (WNN) multimodal integration (Hao et al., 2021; Stuart et al., 2020). Merged multimodal data from the first and last timepoints (day 0 & day 8) contain cells representative of all identified phenotypes.

(B) Distribution of called ATAC peaks per cell by timepoint and phenotype in both donors (***p < 1e-15, one-sided Kolmogorov-Smirnov test).

(C) Overview of global differentially accessible peak (DAP) -linked differentially expressed genes (DEGs) and cluster comparisons for major trajectories of interest. DAPs are identified by their presence in one or more peaks but not in (!) one or more other peaks. DAP-linked DEGs were explored in resting versus arrested cells (c38 vs c47), EBV$^+$ activated versus resting cells (c2 vs c38), and EBV$^+$ activated vs EBV+ differentiated cells (c2 vs c5).

(D) Virus-induced arrest responses. State-resolved DAPs (*c38!c47*) and joint assay trend summaries are presented in addition to example DEGs and DAP linkages resolved by phenotype.

(E) Successful infection trajectory. State-resolved DAPs (*c2!c38*), joint assay summaries, and trajectory-specific examples are presented as in (D).

(F) EBV-induced B cell activation/differentiation continuum. State-resolved DAPs (*c2!c5*), joint assay summaries, and trajectory-specific examples are presented as in (D and E).

See also Figures S26-S38

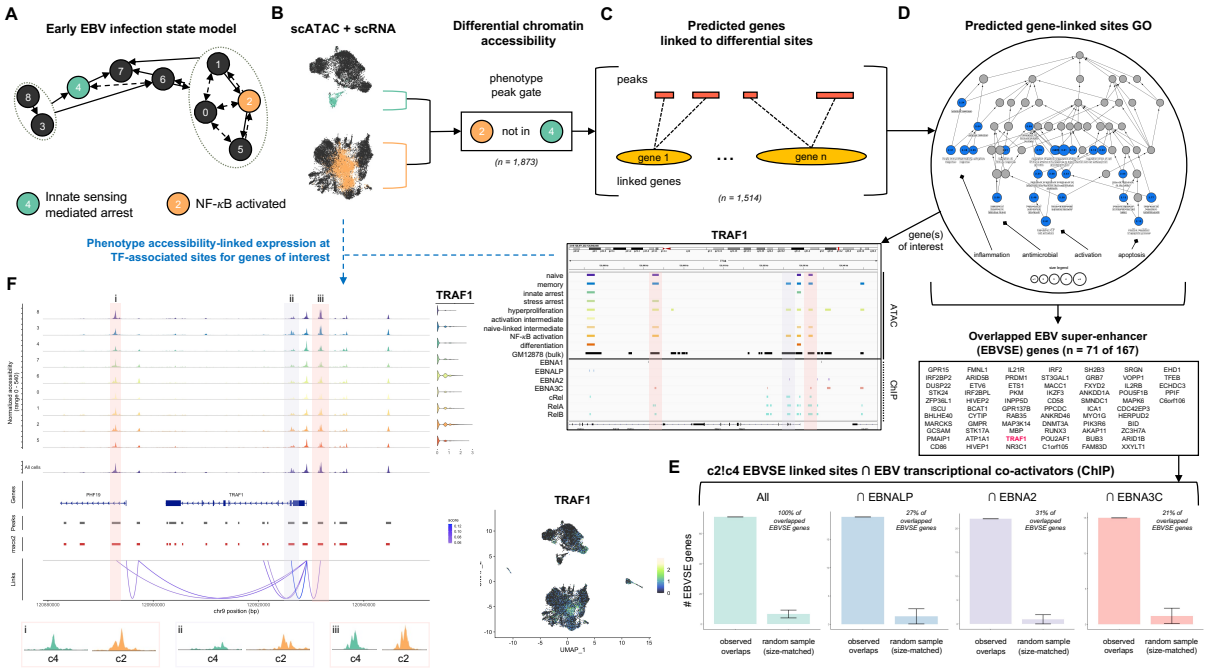

Figure 7. crisp-ATAC analysis of DAP-linked DEGs in activated versus innate arrested EBV⁺ B cells

(A) Schematic of NF-κB activation (c2) and innate arrest (c4) model phenotypes.

(B) Multimodal assay gating to extract *c2!c4* DAPs.

(C and D) Prediction of *cis*-regulatory linkage. All *c2!c4* peak intervals (n=1,873) are used as inputs to the GREAT (McLean et al., 2010) to predict *c2!c4* DAP-linked genes (n=1,514).

(E) Occurrence of EBVSE-linked genes identified as *c2!c4* DAP-linked DEGs associated with select EBNA binding sites relative to expected frequency due to random overlap (n=100 simulation trials, error bars depict mean +/- standard deviation). Random samples were size-matched relative to each EBNA-associated gene list.

(F) Example gene of interest analysis for *TRAF1*, an EBVSE-linked gene identified as a *c2!c4* DAP-linked DEG with phenotype-variable accessibility at multiple EBNA binding sites.

See also Figures S39-S46

875 **References**
876

877 Alfieri, C., Birkenbach, M., and Kieff, E. (1991). Early events in Epstein-Barr virus infection of human B
878 lymphocytes. Virology *181*, 595-608.

879 Allday, M.J., Bazot, Q., and White, R.E. (2015). The EBNA3 family: two oncoproteins and a tumour
880 suppressor that are central to the biology of EBV in B cells. Epstein Barr Virus Volume 2, 61-117.

881 Altmann, M., Pich, D., Ruiss, R., Wang, J., Sugden, B., and Hammerschmidt, W. (2006). Transcriptional
882 activation by EBV nuclear antigen 1 is essential for the expression of EBV's transforming genes.
883 Proceedings of the National Academy of Sciences *103*, 14188-14193.

884 Anderson, L.J., and Longnecker, R. (2008). EBV LMP2A provides a surrogate pre-B cell receptor signal
885 through constitutive activation of the ERK/MAPK pathway. The Journal of general virology *89*, 1563.

886 Arvey, A., Tempera, I., Tsai, K., Chen, H.-S., Tikhmyanova, N., Klichinsky, M., Leslie, C., and Lieberman,
887 P.M. (2012). An atlas of the Epstein-Barr virus transcriptome and epigenome reveals host-virus regulatory
888 interactions. Cell host & microbe *12*, 233-245.

889 Babcock, G.J.D., L.L.; Volk, M.; Thorley-Lawson, D.A. (1998). EBV Persistence in Memory B Cells In
890 Vivo. Immunity *9*, 395-404.

891 Banerjee, S., Lu, J., Cai, Q., Sun, Z., Jha, H.C., and Robertson, E.S. (2014). EBNA3C augments Pim-1
892 mediated phosphorylation and degradation of p21 to promote B-cell proliferation. PLoS Pathog *10*,
893 e1004304.

894 Bartkova, J., Rezaei, N., Liontos, M., Karakaidos, P., Kletsas, D., Issaeva, N., Vassiliou, L.-V.F., Kolettas,
895 E., Niforou, K., and Zoumpourlis, V.C. (2006). Oncogene-induced senescence is part of the
896 tumorigenesis barrier imposed by DNA damage checkpoints. Nature *444*, 633-637.

897 Bird, A.G.M., S.M.; Britton, S. (1981). Cyclosporin A promotes spontaneous outgrowth in vitro of Epstein-
898 Barr virus-induced B-cell lines. Nature *289*, 300-301.

899 Bjornevik, K., Cortese, M., Healy, B.C., Kuhle, J., Mina, M.J., Leng, Y., Elledge, S.J., Niebuhr, D.W.,
900 Scher, A.I., and Munger, K.L. (2022). Longitudinal analysis reveals high prevalence of Epstein-Barr virus
901 associated with multiple sclerosis. Science.

902 Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P., Chang, H.Y., and
903 Greenleaf, W.J. (2015). Single-cell chromatin accessibility reveals principles of regulatory variation.
904 Nature *523*, 486-490.

905 Cahir-McFarland, E.D., Carter, K., Rosenwald, A., Giltnane, J.M., Henrickson, S.E., Staudt, L.M., and
906 Kieff, E. (2004). Role of NF-κB in cell survival and transcription of latent membrane protein 1-expressing
907 or Epstein-Barr virus latency III-infected cells. Journal of virology *78*, 4108-4119.

908 Calender, A., Billaud, M., Aubry, J.-P., Banchereau, J., Vuillaume, M., and Lenoir, G.M. (1987). Epstein-
909 Barr virus (EBV) induces expression of B-cell activation markers on in vitro infection of EBV-negative B-
910 lymphoma cells. Proceedings of the National Academy of Sciences *84*, 8060-8064.

911 Canaan, A., Haviv, I., Urban, A.E., Schulz, V.P., Hartman, S., Zhang, Z., Palejev, D., Deisseroth, A.B.,
912 Lacy, J., and Snyder, M. (2009). EBNA1 regulates cellular gene expression by binding cellular promoters.
913 Proceedings of the National Academy of Sciences *106*, 22421-22426.

914 Cao, J., Cusanovich, D.A., Ramani, V., Aghamirzaie, D., Pliner, H.A., Hill, A.J., Daza, R.M., McFaline-
915 Figueroa, J.L., Packer, J.S., and Christiansen, L. (2018). Joint profiling of chromatin accessibility and
916 gene expression in thousands of single cells. Science *361*, 1380-1385.

917 Chen, S., Lake, B.B., and Zhang, K. (2019). High-throughput sequencing of the transcriptome and
918 chromatin accessibility in the same cell. Nature biotechnology *37*, 1452-1457.

919 Clark, S.J., Argelaguet, R., Kapourani, C.-A., Stubbs, T.M., Lee, H.J., Alda-Catalinas, C., Krueger, F.,
920 Sanguinetti, G., Kelsey, G., and Marioni, J.C. (2018). scNMT-seq enables joint profiling of chromatin
921 accessibility DNA methylation and transcription in single cells. Nature communications *9*, 1-9.

922 Cohen, J.I., Fauci, A.S., Varmus, H., and Nabel, G.J. (2011). Epstein-Barr virus: an important vaccine
923 target for cancer prevention. Science translational medicine *3*, 107fs107-107fs107.

924 Cohen, J.I., Wang, F., Mannick, J., and Kieff, E. (1989). Epstein-Barr virus nuclear protein 2 is a key
925 determinant of lymphocyte transformation. Proceedings of the National Academy of Sciences *86*, 9558-
926 9562.

927 Courtois-Cox, S., Jones, S.L., and Cichowski, K. (2008). Many roads lead to oncogene-induced
928 senescence. Oncogene *27*, 2801-2809.

929 Deisenroth, C., and Zhang, Y. (2010). Ribosome biogenesis surveillance: probing the ribosomal protein-
930 Mdm2-p53 pathway. Oncogene *29*, 4253-4260.

931 Devergne, O., Hatzivassiliou, E., Izumi, K.M., Kaye, K.M., Kleijnen, M.F., Kieff, E., and Mosialos, G.
932 (1996). Association of TRAF1, TRAF2, and TRAF3 with an Epstein-Barr virus LMP1 domain important for
933 B-lymphocyte transformation: role in NF-kappaB activation. Molecular and cellular biology *16*, 7098-7108.

934 Devergne, O., McFarland, E.C., Mosialos, G., Izumi, K.M., Ware, C.F., and Kieff, E. (1998). Role of the
935 TRAF binding site and NF-κB activation in Epstein-Barr virus latent membrane protein 1-induced cell
936 gene expression. Journal of virology *72*, 7900-7908.

937 Dheekollu, J., Wiedmer, A., Ayyanathan, K., Deakyne, J.S., Messick, T.E., and Lieberman, P.M. (2021).
938 Cell-cycle-dependent EBNA1-DNA crosslinking promotes replication termination at oriP and viral episome
939 maintenance. Cell *184*, 643-654. e613.

940 Di Micco, R., Sulli, G., Dobreva, M., Liontos, M., Botrugno, O.A., Gargiulo, G., Dal Zuffo, R., Matti, V.,
941 d'Ario, G., and Montani, E. (2011). Interplay between oncogene-induced DNA damage response and
942 heterochromatin in senescence and cancer. Nature cell biology *13*, 292-302.

943 Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L.
944 (2016). Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. Cell systems
945 *3*, 99-101.

946 Efremova, M., and Teichmann, S.A. (2020). Computational methods for single-cell omics across
947 modalities. Nature methods *17*, 14-17.

948 Eliopoulos, A.G., Waites, E.R., Blake, S.M., Davies, C., Murray, P., and Young, L.S. (2003). TRAF1 is a
949 critical regulator of JNK signaling by the TRAF-binding domain of the Epstein-Barr virus-encoded latent
950 infection membrane protein 1 but not CD40. Journal of virology *77*, 1316-1328.

951 Epstein, M., Achong, B., and Barr, Y. (1964). Virus particles in cultured lymphoblasts from Burkitt's
952 lymphoma. Lancet, 702-703.

953 Fish, K., Comoglio, F., Shaffer, A.L., Ji, Y., Pan, K.-T., Scheich, S., Oellerich, A., Doebele, C., Ikeda, M.,
954 and Schaller, S.J. (2020). Rewiring of B cell receptor signaling by Epstein–Barr virus LMP2A.
955 Proceedings of the National Academy of Sciences *117*, 26318-26327.

956 Forte, E., and Luftig, M.A. (2009). MDM2-dependent inhibition of p53 is required for Epstein-Barr virus B-
957 cell growth transformation and infected-cell survival. Journal of virology *83*, 2491-2499.

958 Fries, K.L., Miller, W.E., and Raab-Traub, N. (1999). The A20 protein interacts with the Epstein–Barr virus
959 latent membrane protein 1 (LMP1) and alters the LMP1/TRAF1/TRADD complex. Virology *264*, 159-166.

960 Gan, L., Seki, A., Shen, K., Iyer, H., Han, K., Hayer, A., Wollman, R., Ge, X., Lin, J.R., and Dey, G.
961 (2019). The lysosomal GPCR-like protein GPR137B regulates Rag and mTORC1 localization and activity.
962 Nature cell biology *21*, 614-626.

963 Gao, J., Xia, L., Lu, M., Zhang, B., Chen, Y., Xu, R., and Wang, L. (2012). TM7SF1 (GPR137B): a novel
964 lysosome integral membrane protein. Molecular biology reports *39*, 8883-8889.

965 Glaros, V., Rauschmeier, R., Artemov, A.V., Reinhardt, A., Ols, S., Emmanouilidi, A., Gustafsson, C.,
966 You, Y., Mirabello, C., and Björklund, Å.K. (2021). Limited access to antigen drives generation of early B
967 cell memory while restraining the plasmablast response. Immunity *54*, 2005-2023. e2010.

968 Glück, S., Guey, B., Gulen, M.F., Wolter, K., Kang, T.-W., Schmacke, N.A., Bridgeman, A., Rehwinkel, J.,
969 Zender, L., and Ablasser, A. (2017). Innate immune sensing of cytosolic chromatin fragments through
970 cGAS promotes senescence. Nature cell biology *19*, 1061-1070.

971 Greenfeld, H., Takasaki, K., Walsh, M.J., Ersing, I., Bernhardt, K., Ma, Y., Fu, B., Ashbaugh, C.W., Cabo,
972 J., and Mollo, S.B. (2015). TRAF1 coordinates polyubiquitin signaling to enhance Epstein-Barr virus
973 LMP1-mediated growth and survival pathway activation. PLoS pathogens *11*, e1004890.

974 Guasparri, I., Bubman, D., and Cesarman, E. (2008). EBV LMP2A affects LMP1-mediated NF-κB
975 signaling and survival of lymphoma cells by regulating TRAF2 expression. Blood, The Journal of the
976 American Society of Hematology *111*, 3813-3820.

977 Guo, R., Jiang, C., Zhang, Y., Govande, A., Trudeau, S.J., Chen, F., Fry, C.J., Puri, R., Wolinsky, E., and
978 Schineller, M. (2020). MYC controls the Epstein-Barr virus lytic switch. Molecular cell *78*, 653-669. e658.

979 Han, I., Harada, S., Weaver, D., Xue, Y., Lane, W., Orstavik, S., Skalhegg, B., and Kieff, E. (2001).
980 EBNA-LP associates with cellular proteins including DNA-PK and HA95. Journal of virology *75*, 2475-
981 2481.

982 Hao, Y., Hao, S., Andersen-Nissen, E., Mauck III, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J.,
983 Darby, C., and Zager, M. (2021). Integrated analysis of multimodal single-cell data. Cell.

984 Harada, S., and Kieff, E. (1997). Epstein-Barr virus nuclear protein LP stimulates EBNA-2 acidic domain-
985 mediated transcriptional activation. Journal of virology *71*, 6611-6618.

986 Henderson, S., Rowe, M., Gregory, C., Croom-Carter, D., Wang, F., Longnecker, R., Kieff, E., and
987 Rickinson, A. (1991). Induction of bcl-2 expression by Epstein-Barr virus latent membrane protein 1
988 protects infected B cells from programmed cell death. Cell *65*, 1107-1115.

989 Henle, W., Diehl, V., Kohn, G., Zur Hausen, H., and Henle, G. (1967). Herpes-type virus and
990 chromosome marker in normal leukocytes after growth with irradiated Burkitt cells. Science *157*, 1064-
991 1065.

992 Hodgkin, P.D., Lee, J.-H., and Lyons, A.B. (1996). B cell differentiation and isotype switching is related to
993 division cycle number. Journal of Experimental Medicine *184*, 277-281.

994 Hoffmann, A., Levchenko, A., Scott, M.L., and Baltimore, D. (2002). The IκB-NF-κB signaling module:
995 temporal control and selective gene activation. science *298*, 1241-1245.

996 Humme, S., Reisbach, G., Feederle, R., Delecluse, H.-J., Bousset, K., Hammerschmidt, W., and
997 Schepers, A. (2003). The EBV nuclear antigen 1 (EBNA1) enhances B cell immortalization several
998 thousandfold. Proceedings of the National Academy of Sciences *100*, 10989-10994.

999 Jiang, S., and Mortazavi, A. (2018). Integrating ChIP-seq with other functional genomics data. Briefings in
1000 functional genomics *17*, 104-115.

1001 Jiang, S., Zhou, H., Liang, J., Gerdt, C., Wang, C., Ke, L., Schmidt, S.C., Narita, Y., Ma, Y., and Wang, S.
1002 (2017). The Epstein-Barr virus regulome in lymphoblastoid cells. Cell host & microbe *22*, 561-573. e564.

1003 Johannsen, E., Luftig, M., Chase, M.R., Weicksel, S., Cahir-McFarland, E., Illanes, D., Sarracino, D., and
1004 Kieff, E. (2004). Proteins of purified Epstein-Barr virus. Proceedings of the National Academy of Sciences
1005 *101*, 16286-16291.

1006 Junker, J.P., and van Oudenaarden, A. (2014). Every cell is special: genome-wide studies add a new
1007 dimension to single-cell biology. Cell *157*, 8-11.

1008 Kilger, E., Kieser, A., Baumann, M., and Hammerschmidt, W. (1998). Epstein–Barr virus-mediated B-cell
1009 proliferation is dependent upon latent membrane protein 1, which simulates an activated CD40 receptor.
1010 The EMBO journal *17*, 1700-1709.

1011 Kim, U., Qin, X.-F., Gong, S., Stevens, S., Luo, Y., Nussenzweig, M., and Roeder, R.G. (1996). The B-
1012 cell-specific transcription coactivator OCA-B/OBF-1/Bob-1 is essential for normal production of
1013 immunoglobulin isotypes. Nature *383*, 542-547.

1014 King, H.W., Orban, N., Riches, J.C., Clear, A.J., Warnes, G., Teichmann, S.A., and James, L.K. (2021).
1015 Single-cell analysis of human B cell maturation predicts how antibody class switching shapes selection
1016 dynamics. Science Immunology *6*, eabe6291.

1017 Krynetskaia, N.F., Phadke, M.S., Jadhav, S.H., and Krynetskiy, E.Y. (2009). Chromatin-associated
1018 proteins HMGB1/2 and PDIA3 trigger cellular response to chemotherapy-induced DNA damage.
1019 Molecular cancer therapeutics *8*, 864-872.

1020 Kwak, I.H., Kim, H.S., Choi, O.R., Ryu, M.S., and Lim, I.K. (2004). Nuclear accumulation of globular actin
1021 as a cellular senescence marker. Cancer research *64*, 572-580.

1022 Laichalk, L.L., and Thorley-Lawson, D.A. (2005). Terminal differentiation into plasma cells initiates the
1023 replicative cycle of Epstein-Barr virus in vivo. J Virol *79*, 1296-1307.

1024 Lamontagne, R.J., Soldan, S.S., Su, C., Wiedmer, A., Won, K.J., Lu, F., Goldman, A.R., Wickramasinghe,
1025 J., Tang, H.-Y., Speicher, D.W.*, et al.* (2021). A multi-omics approach to Epstein-Barr virus
1026 immortalization of B-cells reveals EBNA1 chromatin pioneering activities targeting nucleotide metabolism.
1027 PLOS Pathogens *17*, e1009208.

1028 Lanz, T.V., Brewer, R.C., Ho, P.P., Moon, J.-S., Jude, K.M., Fernandez, D., Fernandes, R.A., Gomez,
1029 A.M., Nadj, G.-S., Bartley, C.M.*, et al.* (2022). Clonally Expanded B Cells in Multiple Sclerosis Bind EBV
1030 EBNA1 and GlialCAM. Nature.

1031 Lenain, C., de Graaf, C.A., Pagie, L., Visser, N.L., de Haas, M., de Vries, S.S., Peric-Hupkes, D., van
1032 Steensel, B., and Peeper, D.S. (2017). Massive reshaping of genome–nuclear lamina interactions during
1033 oncogene-induced senescence. Genome research *27*, 1634-1644.

1034 Lessard, F., Igelmann, S., Trahan, C., Huot, G., Saint-Germain, E., Mignacca, L., Del Toro, N., Lopes-
1035 Paciencia, S., Le Calvé, B., and Montero, M. (2018). Senescence-associated ribosome biogenesis
1036 defects contributes to cell cycle arrest through the Rb pathway. Nature cell biology *20*, 789-799.

1037 Lindahl, T., Adams, A., Bjursell, G., Bornkamm, G.W., Kaschka-Dierich, C., and Jehn, U. (1976).
1038 Covalently closed circular duplex DNA of Epstein-Barr virus in a human lymphoid cell line. Journal of
1039 molecular biology *102*, 511-530.

1040 Ling, P.D., Peng, R.S., Nakajima, A., Yu, J.H., Tan, J., Moses, S.M., Yang, W.H., Zhao, B., Kieff, E., and
1041 Bloch, K.D. (2005). Mediation of Epstein–Barr virus EBNA-LP transcriptional coactivation by Sp100. The
1042 EMBO journal *24*, 3565-3575.

1043 Liu, T. (2014). Use model-based analysis of ChIP-Seq (MACS) to analyze short reads generated by
1044 sequencing protein–DNA interactions in embryonic stem cells. In Stem Cell Transcriptional Networks
1045 (Springer), pp. 81-95.

1046 Longnecker, R.M., Kieff, E., and Cohen, J.I. (2013). Epstein-barr virus. In Fields Virology: Sixth Edition
1047 (Wolters Kluwer Health Adis (ESP)).

1048 Lu, F., Chen, H.-S., Kossenkov, A.V., DeWispeleare, K., Won, K.-J., and Lieberman, P.M. (2016). EBNA2
1049 drives formation of new chromosome binding sites and target genes for B-cell master regulatory
1050 transcription factors RBP-jκ and EBF1. PLoS pathogens *12*, e1005339.

1051 Lu, F., Wikramasinghe, P., Norseen, J., Tsai, K., Wang, P., Showe, L., Davuluri, R.V., and Lieberman,
1052 P.M. (2010). Genome-wide analysis of host-chromosome binding sites for Epstein-Barr Virus Nuclear
1053 Antigen 1 (EBNA1). Virology journal *7*, 1-17.

1054 Luftig, M., Prinarakis, E., Yasui, T., Tsichritzis, T., Cahir-McFarland, E., Inoue, J.-I., Nakano, H., Mak,
1055 T.W., Yeh, W.-C., and Li, X. (2003). Epstein–Barr virus latent membrane protein 1 activation of NF-κB
1056 through IRAK1 and TRAF6. Proceedings of the National Academy of Sciences *100*, 15595-15600.

1057 Lünemann, A., Rowe, M., and Nadal, D. (2015). Innate immune recognition of EBV. Epstein Barr Virus
1058 Volume 2, 265-287.

1059 Luo, Y., and Roeder, R.G. (1995). Cloning, functional characterization, and mechanism of action of the B-
1060 cell-specific transcriptional coactivator OCA-B. Molecular and cellular biology *15*, 4115-4124.

1061 Lupton, S., and Levine, A.J. (1985). Mapping genetic elements of Epstein-Barr virus that facilitate
1062 extrachromosomal persistence of Epstein-Barr virus-derived plasmids in human cells. Molecular and
1063 Cellular Biology *5*, 2533-2542.

1064 Ma, S., Zhang, B., LaFave, L.M., Earl, A.S., Chiang, Z., Hu, Y., Ding, J., Brack, A., Kartha, V.K., and Tay,
1065 T. (2020). Chromatin potential identified by shared single-cell profiling of RNA and chromatin. Cell *183*,
1066 1103-1116. e1120.

1067 Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R.,
1068 Kamitaki, N., and Martersteck, E.M. (2015). Highly parallel genome-wide expression profiling of individual
1069 cells using nanoliter droplets. Cell *161*, 1202-1214.

1070 Mannick, J., Cohen, J., Birkenbach, M., Marchini, A., and Kieff, E. (1991). The Epstein-Barr virus nuclear
1071 protein encoded by the leader of the EBNA RNAs is important in B-lymphocyte transformation. Journal of
1072 virology *65*, 6826-6837.

1073 Martin, H.J., Lee, J.M., Walls, D., and Hayward, S.D. (2007). Manipulation of the toll-like receptor 7
1074 signaling pathway by Epstein-Barr virus. Journal of virology *81*, 9748-9758.

1075 Martina, J.A., Chen, Y., Gucek, M., and Puertollano, R. (2012). MTORC1 functions as a transcriptional
1076 regulator of autophagy by preventing nuclear transport of TFEB. Autophagy *8*, 903-914.

1077 Maruo, S., Zhao, B., Johannsen, E., Kieff, E., Zou, J., and Takada, K. (2011). Epstein-Barr virus nuclear
1078 antigens 3C and 3A maintain lymphoblastoid cell growth by repressing p16INK4A and p14ARF
1079 expression. Proceedings of the National Academy of Sciences *108*, 1919-1924.

1080 Matsuda, G., Nakajima, K., Kawaguchi, Y., Yamanashi, Y., and Hirai, K. (2003). Epstein-Barr virus (EBV)
1081 nuclear antigen leader protein (EBNA-LP) forms complexes with a cellular anti-apoptosis protein Bcl-2 or
1082 its EBV counterpart BHRF1 through HS1-associated protein X-1. Microbiology and immunology *47*, 91-
1083 99.

1084 McClellan, M.J., Wood, C.D., Ojeniyi, O., Cooper, T.J., Kanhere, A., Arvey, A., Webb, H.M., Palermo,
1085 R.D., Harth-Hertle, M.L., Kempkes, B.*, et al.* (2013). Modulation of enhancer looping and differential gene
1086 targeting by Epstein-Barr virus transcription factors directs cellular reprogramming. PLoS Pathog *9*,
1087 e1003636.

1088 McFadden, K., Hafez, A.Y., Kishton, R., Messinger, J.E., Nikitin, P.A., Rathmell, J.C., and Luftig, M.A.
1089 (2016). Metabolic stress is a barrier to Epstein–Barr virus-mediated B-cell immortalization. Proceedings of
1090 the National Academy of Sciences *113*, E782-E790.

1091 McInnes, L., Healy, J., and Melville, J. (2018). Umap: Uniform manifold approximation and projection for
1092 dimension reduction. arXiv preprint arXiv:180203426.

1093 McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano,
1094 G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nature biotechnology *28*,
1095 495-501.

1096 Messinger, J.E., Dai, J., Stanland, L.J., Price, A.M., and Luftig, M.A. (2019). Identification of Host
1097 Biomarkers of Epstein-Barr Virus Latency IIb and Latency III. mBio *10*.

1098 Minamitani, T., Yasui, T., Ma, Y., Zhou, H., Okuzaki, D., Tsai, C.-Y., Sakakibara, S., Gewurz, B.E., Kieff,
1099 E., and Kikutani, H. (2015). Evasion of affinity-based selection in germinal centers by Epstein–Barr virus
1100 LMP2A. Proceedings of the National Academy of Sciences *112*, 11612-11617.

1101 Mitchell, S., Roy, K., Zangle, T.A., and Hoffmann, A. (2018). Nongenetic origins of cell-to-cell variability in
1102 B lymphocyte proliferation. Proc Natl Acad Sci U S A *115*, E2888-E2897.

1103 Miyashita, E.M., Yang, B., Babcock, G.J., and Thorley-Lawson, D.A. (1997). Identification of the site of
1104 Epstein-Barr virus persistence in vivo as a resting B cell. Journal of virology *71*, 4882-4891.

1105 Mrozek-Gorska, P., Buschle, A., Pich, D., Schwarzmayr, T., Fechtner, R., Scialdone, A., and
1106 Hammerschmidt, W. (2019). Epstein-Barr virus reprograms human B lymphocytes immediately in the
1107 prelatent phase of infection. Proc Natl Acad Sci U S A *116*, 16046-16055.

1108 Nagelkerke, A., Sieuwerts, A.M., Bussink, J., Sweep, F., Look, M.P., Foekens, J.A., Martens, J., and
1109 Span, P.N. (2014). LAMP3 is involved in tamoxifen resistance in breast cancer cells through the
1110 modulation of autophagy. Endocr Relat Cancer *21*, 101-112.

1111 Nikitin, P.A., Yan, C.M., Forte, E., Bocedi, A., Tourigny, J.P., White, R.E., Allday, M.J., Patel, A., Dave,
1112 S.S., and Kim, W. (2010). An ATM/Chk2-mediated DNA damage-responsive signaling pathway
1113 suppresses Epstein-Barr virus transformation of primary human B cells. Cell host & microbe *8*, 510-522.

1114 Nilsson, K., Klein, G., Henle, W., and Henle, G. (1971). The establishment of lymphoblastoid lines from
1115 adult and fetal human lymphoid tissue and its dependence on EBV. International Journal of Cancer *8*,
1116 443-450.

1117 Nishimura, K., Kumazawa, T., Kuroda, T., Katagiri, N., Tsuchiya, M., Goto, N., Furumai, R., Murayama,
1118 A., Yanagisawa, J., and Kimura, K. (2015). Perturbation of ribosome biogenesis drives cells into
1119 senescence through 5S RNP-mediated p53 activation. Cell reports *10*, 1310-1323.

1120 Nonoyama, M., and Pagano, J.S. (1972). Separation of Epstein-Barr virus DNA from large chromosomal
1121 DNA in non-virus-producing cells. Nature New Biology *238*, 169-171.

1122 O'Dea, E.L., Barken, D., Peralta, R.Q., Tran, K.T., Werner, S.L., Kearns, J.D., Levchenko, A., and
1123 Hoffmann, A. (2007). A homeostatic model of IkappaB metabolism to control constitutive NF-kappaB
1124 activity. Mol Syst Biol *3*, 111.

1125 Osorio, D., and Cai, J.J. (2021). Systematic determination of the mitochondrial proportion in human and
1126 mice tissues for single-cell RNA-sequencing data quality control. Bioinformatics *37*, 963-967.

1127 Osorio, D., Yu, X., Yu, P., Serpedin, E., and Cai, J.J. (2019). Single-cell RNA sequencing of a European
1128 and an African lymphoblastoid cell line. Sci Data *6*, 112.

1129 Osorio, D., Yu, X., Zhong, Y., Li, G., Serpedin, E., Huang, J.Z., and Cai, J.J. (2020). Single-cell
1130 expression variability implies cell function. Cells *9*, 14.

1131 Parker, G.A., Crook, T., Bain, M., Sara, E.A., Farrell, P.J., and Allday, M.J. (1996). Epstein-Barr virus
1132 nuclear antigen (EBNA) 3C is an immortalizing oncoprotein with similar properties to adenovirus E1A and
1133 papillomavirus E7. Oncogene *13*, 2541-2549.

1134 Paschos, K., Parker, G.A., Watanatanasup, E., White, R.E., and Allday, M.J. (2012). BIM promoter
1135 directly targeted by EBNA3C in polycomb-mediated repression by EBV. Nucleic acids research *40*, 7233-
1136 7246.

1137 Peng, R., Moses, S.C., Tan, J., Kremmer, E., and Ling, P.D. (2005). The Epstein-Barr virus EBNA-LP
1138 protein preferentially coactivates EBNA2-mediated stimulation of latent membrane proteins expressed
1139 from the viral divergent promoter. Journal of virology *79*, 4492-4505.

1140 Pich, D., Mrozek-Gorska, P., Bouvet, M., Sugimoto, A., Akidil, E., Grundhoff, A., Hamperl, S., Ling, P.D.,
1141 and Hammerschmidt, W. (2019). First days in the life of naive human B lymphocytes infected with
1142 Epstein-Barr virus. MBio *10*, e01723-01719.

1143 Pope, J., Horne, M., and Scott, W. (1968). Transformation of foetal human leukocytes in vitro by filtrates
1144 of a human leukaemic cell line containing herpes-like virus. International journal of cancer *3*, 857-866.

1145 Portis, T., and Longnecker, R. (2004). Epstein–Barr virus (EBV) LMP2A mediates B-lymphocyte survival
1146 through constitutive activation of the Ras/PI3K/Akt pathway. Oncogene *23*, 8619-8628.

1147 Price, A.M., Dai, J., Bazot, Q., Patel, L., Nikitin, P.A., Djavadian, R., Winter, P.S., Salinas, C.A., Barry,
1148 A.P., Wood, K.C.*, et al.* (2017). Epstein-Barr virus ensures B cell survival by uniquely modulating
1149 apoptosis at early and late times after infection. Elife *6*.

1150 Price, A.M., and Luftig, M.A. (2015). To be or not IIb: a multi-step process for Epstein-Barr virus latency
1151 establishment and consequences for B cell tumorigenesis. PLoS Pathog *11*, e1004656.

1152 Qadir, F.S., Saad; Dominguez-Bendala, Juan (2019). 3D Plotting of scRNAseq data using Seurat objects
1153 (1.3) (Zenodo).

1154 Qadir, M.M.F., Álvarez-Cubela, S., Klein, D., van Dijk, J., Muñiz-Anquela, R., Moreno-Hernández, Y.B.,
1155 Lanzoni, G., Sadiq, S., Navarro-Rubio, B., and García, M.T. (2020). Single-cell resolution analysis of the
1156 human pancreatic ductal progenitor cell niche. Proceedings of the National Academy of Sciences *117*,
1157 10876-10887.

1158 Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017). Reversed graph
1159 embedding resolves complex single-cell trajectories. Nature methods *14*, 979.

1160 Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features.
1161 Bioinformatics *26*, 841-842.

1162 Raab-Traub, N. (2007). EBV-induced oncogenesis. Human herpesviruses: biology, therapy, and
1163 immunoprophylaxis.

1164 Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y., and Tyagi, S. (2006). Stochastic mRNA synthesis in
1165 mammalian cells. PLoS Biol *4*, e309.

1166 Raj, A., Rifkin, S.A., Andersen, E., and Van Oudenaarden, A. (2010). Variability in gene expression
1167 underlies incomplete penetrance. Nature *463*, 913-918.

1168 Raj, A., and Van Oudenaarden, A. (2008). Nature, nurture, or chance: stochastic gene expression and its
1169 consequences. Cell *135*, 216-226.

1170 Ramírez, F., Bhardwaj, V., Arrigoni, L., Lam, K.C., Grüning, B.A., Villaveces, J., Habermann, B., Akhtar,
1171 A., and Manke, T. (2018). High-resolution TADs reveal DNA sequences underlying genome organization
1172 in flies. Nature communications *9*, 1-15.

1173 Ressing, M.E., van Gent, M., Gram, A.M., Hooykaas, M.J., Piersma, S.J., and Wiertz, E.J. (2015).
1174 Immune evasion by Epstein-Barr virus. Epstein Barr Virus Volume 2, 355-381.

1175 Rickinson, A., and Kieff, E. (2007). Epstein-Barr virus, p 2655–2700. Fields virology *2*.

1176 Robertson, E.S., Grossman, S., Johannsen, E., Miller, C., Lin, J., Tomkinson, B., and Kieff, E. (1995).
1177 Epstein-Barr virus nuclear protein 3C modulates transcription through interaction with the sequence-
1178 specific DNA-binding protein J kappa. Journal of virology *69*, 3108-3116.

1179 Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P.
1180 (2011). Integrative genomics viewer. Nature biotechnology *29*, 24-26.

1181 Roy, K., Mitchell, S., Liu, Y., Ohta, S., Lin, Y.-s., Metzig, M.O., Nutt, S.L., and Hoffmann, A. (2019). A
1182 regulatory circuit controlling the dynamics of NFκB cRel transitions B cells from proliferation to plasma cell
1183 differentiation. Immunity *50*, 616-628. e616.

1184 Saha, A., Jha, H.C., Upadhyay, S.K., and Robertson, E.S. (2015). Epigenetic silencing of tumor
1185 suppressor genes during in vitro Epstein–Barr virus infection. Proceedings of the National Academy of
1186 Sciences *112*, E5199-E5207.

1187 Sandberg, M., Hammerschmidt, W., and Sugden, B. (1997). Characterization of LMP-1's association with
1188 TRAF1, TRAF2, and TRAF3. Journal of Virology *71*, 4649-4656.

1189 Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-
1190 cell gene expression data. Nature biotechnology *33*, 495-502.

1191 Satpathy, A.T., Granja, J.M., Yost, K.E., Qi, Y., Meschi, F., McDermott, G.P., Olsen, B.N., Mumbach,
1192 M.R., Pierce, S.E., and Corces, M.R. (2019). Massively parallel single-cell chromatin landscapes of
1193 human immune cell development and intratumoral T cell exhaustion. Nature biotechnology *37*, 925-936.

1194 Schubart, D.B., Rolink, A., Kosco-Vilbois, M.H., Botteri, F., and Matthias, P. (1996). B-cell-specif ic
1195 coactivator OBF-1/OCA-B/Bob1 required for immune response and germinal centre formation. Nature
1196 *383*, 538-542.

1197 Settembre, C., Di Malta, C., Polito, V.A., Arencibia, M.G., Vetrini, F., Erdin, S., Erdin, S.U., Huynh, T.,
1198 Medina, D., and Colella, P. (2011). TFEB links autophagy to lysosomal biogenesis. science *332*, 1429-
1199 1433.

1200 Shair, K.H., Schnegg, C.I., and Raab-Traub, N. (2008). EBV latent membrane protein 1 effects on
1201 plakoglobin, cell growth, and migration. Cancer research *68*, 6997-7005.

1202 Shalek, A.K., Satija, R., Adiconis, X., Gertner, R.S., Gaublomme, J.T., Raychowdhury, R., Schwartz, S.,
1203 Yosef, N., Malboeuf, C., and Lu, D. (2013). Single-cell transcriptomics reveals bimodality in expression
1204 and splicing in immune cells. Nature *498*, 236-240.

1205 Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and
1206 Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction
1207 networks. Genome research *13*, 2498-2504.

1208 Shapiro, E., Biezuner, T., and Linnarsson, S. (2013). Single-cell sequencing-based technologies will
1209 revolutionize whole-organism science. Nature Reviews Genetics *14*, 618-630.

1210 Simpson, E.H. (1951). The interpretation of interaction in contingency tables. Journal of the Royal
1211 Statistical Society: Series B (Methodological) *13*, 238-241.

1212 Sinclair, A.J., Palmero, I., Peters, G., and Farrell, P. (1994). EBNA-2 and EBNA-LP cooperate to cause
1213 G0 to G1 transition during immortalization of resting human B lymphocytes by Epstein-Barr virus. The
1214 EMBO journal *13*, 3321-3328.

1215 Skalska, L., White, R.E., Franz, M., Ruhmann, M., and Allday, M.J. (2010). Epigenetic repression of
1216 p16INK4A by latent Epstein-Barr virus requires the interaction of EBNA3A and EBNA3C with CtBP. PLoS
1217 pathogens *6*, e1000951.

1218 Smith, N., Tierney, R., Wei, W., Vockerodt, M., Murray, P.G., Woodman, C.B., and Rowe, M. (2013).
1219 Induction of interferon-stimulated genes on the IL-4 response axis by Epstein-Barr virus infected human b
1220 cells; relevance to cellular transformation. PLoS One *8*, e64868.

1221 SoRelle, E.D., Dai, J., Bonglack, E.N., Heckenberg, E.M., Zhou, J.Y., Giamberardino, S.N., Bailey, J.A.,
1222 Gregory, S.G., Chan, C., and Luftig, M.A. (2021). Single-cell RNA-seq reveals transcriptomic
1223 heterogeneity mediated by host–pathogen dynamics in lymphoblastoid cell lines. Elife *10*, e62586.

1224 Stros, M., Ozaki, T., Bacıková, A., Kageyama, H., and Nakagawara, A. (2002). HMGB1 and HMGB2 cell-
1225 specifically down-regulate the p53-and p73-dependent sequence-specific transactivation from the human
1226 Bax gene promoter. Journal of Biological Chemistry *277*, 7157-7164.

1227 Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck III, W.M., Hao, Y., Stoeckius, M.,
1228 Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. Cell *177*, 1888-1902.
1229 e1821.

1230 Stuart, T., Srivastava, A., Lareau, C., and Satija, R. (2020). Multimodal single-cell chromatin analysis with
1231 Signac. BioRxiv.

1232 Stuart, T., Srivastava, A., Madad, S., Lareau, C.A., and Satija, R. (2021). Single-cell chromatin state
1233 analysis with Signac. Nature Methods, 1-9.

1234 Styles, C.T., Bazot, Q., Parker, G.A., White, R.E., Paschos, K., and Allday, M.J. (2017). EBV
1235 epigenetically suppresses the B cell-to-plasma cell differentiation pathway while establishing long-term
1236 latency. PLoS Biol *15*, e2001992.

1237 Suan, D., Kräutler, N.J., Maag, J.L., Butt, D., Bourne, K., Hermes, J.R., Avery, D.T., Young, C., Statham,
1238 A., and Elliott, M. (2017). CCR6 defines memory B cell precursors in mouse and human germinal centers,
1239 revealing light-zone location and predominant low antigen affinity. Immunity *47*, 1142-1153. e1144.

1240 Sun, C.C., and Thorley-Lawson, D.A. (2007). Plasma cell-specific transcription factor XBP-1s binds to
1241 and transactivates the Epstein-Barr virus BZLF1 promoter. J Virol *81*, 13566-13577.

1242 Szymula, A., Palermo, R.D., Bayoumy, A., Groves, I.J., Ba Abdullah, M., Holder, B., and White, R.E.
1243 (2018). Epstein-Barr virus nuclear antigen EBNA-LP is essential for transforming naive B cells, and
1244 facilitates recruitment of transcription factors to the viral genome. PLoS pathogens *14*, e1006890.

1245 Taylor, G.M., Raghuwanshi, S.K., Rowe, D.T., Wadowsky, R.M., and Rosendorff, A. (2011). Endoplasmic
1246 reticulum stress causes EBV lytic replication. Blood, The Journal of the American Society of Hematology
1247 *118*, 5528-5539.

1248 Taylor, J.J., Pape, K.A., Steach, H.R., and Jenkins, M.K. (2015). Apoptosis and antigen affinity limit
1249 effector cell differentiation of a single naïve B cell. Science *347*, 784-787.

1250 Thorley-Lawson, D.A. (2001). Epstein-Barr virus: exploiting the immune system. Nature Reviews
1251 Immunology *1*, 75-82.

1252 Thorley-Lawson, D.A., and Gross, A. (2004). Persistence of the Epstein–Barr virus and the origins of
1253 associated lymphomas. New England Journal of Medicine *350*, 1328-1337.

1254 Thorley-Lawson, D.A., and Mann, K.P. (1985). Early events in Epstein-Barr virus infection provide a
1255 model for B cell activation. The Journal of experimental medicine *162*, 45-59.

1256 Tomkinson, B., Robertson, E., and Kieff, E. (1993). Epstein-Barr virus nuclear proteins EBNA-3A and
1257 EBNA-3C are essential for B-lymphocyte growth transformation. Journal of virology *67*, 2014-2025.

1258 Tosato, G., Blaese, R., and Yarchoan, R. (1985). Relationship between immunoglobulin production and
1259 immortalization by Epstein Barr virus. The Journal of Immunology *135*, 959-964.

1260 Trapnell, C. (2015). Defining cell types and states with single-cell genomics. Genome research *25*, 1491-
1261 1498.

1262 Tsai, K., Thikmyanova, N., Wojcechowskyj, J.A., Delecluse, H.-J., and Lieberman, P.M. (2011). EBV
1263 tegument protein BNRF1 disrupts DAXX-ATRX to activate viral early gene transcription. PLoS pathogens
1264 *7*, e1002376.

1265 Uchida, J., Yasui, T., Takaoka-Shichijo, Y., Muraoka, M., Kulwichit, W., Raab-Traub, N., and Kikutani, H.
1266 (1999). Mimicry of CD40 signals by Epstein-Barr virus LMP1 in B lymphocyte responses. Science *286*,
1267 300-303.

1268 Wang, A., Welch, R., Zhao, B., Ta, T., Keleş, S., and Johannsen, E. (2015). Epstein-Barr virus nuclear
1269 antigen 3 (EBNA3) proteins regulate EBNA2 binding to distinct RBPJ genomic sites. Journal of virology
1270 *90*, 2906-2919.

1271 Wang, C., Li, D., Zhang, L., Jiang, S., Liang, J., Narita, Y., Hou, I., Zhong, Q., Zheng, Z., and Xiao, H.
1272 (2019). RNA sequencing analyses of gene expression during Epstein-Barr virus infection of primary B
1273 lymphocytes. Journal of virology *93*, e00226-00219.

1274 Wang, L.W., Jiang, S., and Gewurz, B.E. (2017). Epstein-Barr virus LMP1-mediated oncogenicity. Journal
1275 of virology *91*, e01718-01716.

1276 Wendel-Hansen, V., Rosén, A., and Klein, G. (1987). EBV-transformed lymphoblastoid cell lines down-
1277 regulate ebna in parallel with secretory differentiation. International journal of cancer *39*, 404-408.

1278 White, R.E., Groves, I.J., Turro, E., Yee, J., Kremmer, E., and Allday, M.J. (2010). Extensive co-operation
1279 between the Epstein-Barr virus EBNA3 proteins in the manipulation of host gene expression and
1280 epigenetic chromatin modification. PloS one *5*, e13979.

1281 Wills, Q.F., Livak, K.J., Tipping, A.J., Enver, T., Goldson, A.J., Sexton, D.W., and Holmes, C. (2013).
1282 Single-cell gene expression analysis reveals genetic associations masked in whole-tissue experiments.
1283 Nature biotechnology *31*, 748-752.

1284 Wolff, J., Bhardwaj, V., Nothjunge, S., Richard, G., Renschler, G., Gilsbach, R., Manke, T., Backofen, R.,
1285 Ramírez, F., and Grüning, B.A. (2018). Galaxy HiCExplorer: a web server for reproducible Hi-C data
1286 analysis, quality control and visualization. Nucleic acids research *46*, W11-W16.

1287 Wolff, J., Rabbani, L., Gilsbach, R., Richard, G., Manke, T., Backofen, R., and Grüning, B.A. (2020).
1288 Galaxy HiCExplorer 3: a web server for reproducible Hi-C, capture Hi-C and single-cell Hi-C data
1289 analysis, quality control and visualization. Nucleic acids research *48*, W177-W184.

1290 Wood, V., O'neil, J., Wei, W., Stewart, S., Dawson, C., and Young, L. (2007). Epstein–Barr virus-encoded
1291 EBNA1 regulates cellular gene transcription and modulates the STAT1 and TGF β signaling pathways.
1292 Oncogene *26*, 4135-4147.

1293 Wu, J., Xia, L., Yao, X., Yu, X., Tumas, K.C., Sun, W., Cheng, Y., He, X., Peng, Y.-c., and Singh, B.K.
1294 (2020). The E3 ubiquitin ligase MARCH1 regulates antimalaria immunity through interferon signaling and
1295 T cell activation. Proceedings of the National Academy of Sciences *117*, 16567-16578.

1296 Wu, Y., Bressette, D., Carrell, J.A., Kaufman, T., Feng, P., Taylor, K., Gan, Y., Cho, Y.H., Garcia, A.D.,
1297 and Gollatz, E. (2000). Tumor necrosis factor (TNF) receptor superfamily member TACI is a high affinity
1298 receptor for TNF family members APRIL and BLyS. Journal of Biological Chemistry *275*, 35478-35485.

1299 Yanai, H., Ban, T., Wang, Z., Choi, M.K., Kawamura, T., Negishi, H., Nakasato, M., Lu, Y., Hangai, S.,
1300 and Koshiba, R. (2009). HMGB proteins function as universal sentinels for nucleic-acid-mediated innate
1301 immune responses. Nature *462*, 99-103.

1302 Yates, J.L., Warren, N., and Sugden, B. (1985). Stable replication of plasmids derived from Epstein–Barr
1303 virus in various mammalian cells. Nature *313*, 812-815.

1304 Yetming, K.D., Lupey-Green, L.N., Biryukov, S., Hughes, D.J., Marendy, E.M., Miranda, J.L., and Sample,
1305 J.T. (2020). The BHLF1 locus of Epstein-Barr virus contributes to viral latency and B-cell immortalization.
1306 Journal of virology *94*, e01215-01220.

1307 Young, L.S., Arrand, J.R., and Murray, P.G. (2007). EBV gene expression and regulation. Human
1308 herpesviruses: biology, therapy, and immunoprophylaxis.

1309 Young, L.S., and Rickinson, A.B. (2004). Epstein-Barr virus: 40 years on. Nat Rev Cancer *4*, 757-768.

1310 Zhao, B., Zou, J., Wang, H., Johannsen, E., Peng, C.-w., Quackenbush, J., Mar, J.C., Morton, C.C.,
1311 Freedman, M.L., and Blacklow, S.C. (2011). Epstein-Barr virus exploits intrinsic B-lymphocyte
1312 transcription programs to achieve immortal cell growth. Proceedings of the National Academy of Sciences
1313 *108*, 14902-14907.

1314 Zheng, G.X., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D.,
1315 McDermott, G.P., and Zhu, J. (2017). Massively parallel digital transcriptional profiling of single cells.
1316 Nature communications *8*, 1-12.

1317 Zhou, H., Schmidt, S.C., Jiang, S., Willox, B., Bernhardt, K., Liang, J., Johannsen, E.C., Kharchenko, P.,
1318 Gewurz, B.E., and Kieff, E. (2015). Epstein-Barr virus oncoprotein super-enhancers control B cell growth.
1319 Cell host & microbe *17*, 205-216.

1320 Zhu, C., Yu, M., Huang, H., Juric, I., Abnousi, A., Hu, R., Lucero, J., Behrens, M.M., Hu, M., and Ren, B.
1321 (2019). An ultra high-throughput method for single-cell joint analysis of open chromatin and
1322 transcriptome. Nature structural & molecular biology *26*, 1063-1070.

1323