1 Title: A three-dimensional Air-Liquid Interface Culture Model for the Study of Epstein-
2 Barr virus Infection in the Nasopharynx
3
4 Running title: 3-D model of EBV infection in the nasopharynx
5
6 Authors: Phillip Ziegler[1], Yulong Bai[2], Yarong Tian[3], Sanna Abrahamsson[3], Anthony
7 Green[4], John Moore[5], Stella E. Lee[5], Michael M. Myerburg[6], Hyun Jung Park[2], Ka-Wei
8 Tang[3], Kathy H.Y. Shair*[1,7]
9
10 These authors contributed equally: YuLong Bai and Yarong Tian.
11
12 Affiliations:
13 [1]Cancer Virology Program, University of Pittsburgh Medical Center (UPMC);
14 [2]Department of Human Genetics; [4]University of Pittsburgh Research Histology Services;
15 [5]UPMC Department of Otolaryngology; [6]Division of Pulmonary, Allergy, and Critical
16 Care Medicine; [7]Department of Microbiology and Molecular Genetics, University of
17 Pittsburgh, Pittsburgh, Pennsylvania, USA.
18
19 [3]Wallenberg Centre for Molecular and Translational Medicine, Sahlgrenska Center for
20 Cancer Research, Department of Infectious Diseases, Institute of Biomedicine,
21 Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden.
22
23 *Contact information for corresponding author: kas361@pitt.edu

ABSTRACT

Epstein-Barr virus (EBV) infection is ubiquitous in humans and is associated with the cancer, nasopharyngeal carcinoma. EBV replicates in the differentiated layers of stratified keratinocytes but whether the other cell types of the airway epithelium are susceptible to EBV is unknown. Here, we demonstrate with primary nasopharyngeal cells grown at the air-liquid interface that the pseudostratified epithelium can be susceptible to EBV infection and we report that susceptible cell types with distinct EBV transcription profiles can be identified by single-cell RNA-sequencing. Although EBV infection in the nasopharynx has evaded detection in asymptomatic carriers, these findings demonstrate that EBV latent and lytic infection can occur in the cells of the nasopharyngeal epithelium.

INTRODUCTION

Epstein-Barr virus (EBV) is a human tumor virus from the γ-herpesvirus family (*1*). Infection is chronic and mostly asymptomatic but in a subset of individuals, latent infection is associated with certain types of B-cell lymphomas and epithelial carcinomas (*1*). EBV-associated nasopharyngeal carcinoma (NPC) is endemic in Southeast Asians and Alaskan Inuits (*2*). Diet and host genetics are risk-factors for NPC but almost all NPC tumors share the characteristic of latent and clonal infection with EBV (*2*). Thus, EBV infection is not a passenger infection but precedes the expansion of the neoplastic cell. EBV immortalizes B-cells; however, there are no reports of immortalization in epithelial cells (*3, 4*). Accordingly, the molecular pathogenesis of EBV in epithelial cells has been enigmatic. In the absence of cancer, EBV-infected cells are rarely detected in

47 the nasopharyngeal epithelium (*5, 6*). This infrequency may be due to robust immune
48 surveillance and/or small areas of infection that are difficult to capture at any one time.
49 Thus, studies on EBV molecular pathogenesis in the nasopharynx have relied on cell
50 culture models. Conventionally, two-dimensional (2-D) culture is used to study EBV
51 latent infection in epithelial cells but it does not recapitulate the many cell types of the
52 three-dimensional (3-D) epithelium in the nasopharynx (*3*). Furthermore, EBV-infected
53 cell lines in 2-D culture are largely refractory to reactivation (*3*). Both latent and lytic
54 infection are thought to encourage the persistence and spread of EBV in the
55 nasopharynx, which can predispose cells to neoplasia (*2*).
56
57 Differentiation-induced reactivation in oral stratified keratinocytes cultured in 3-D
58 organotypic rafts explains the lytic pathology of EBV-associated oral hairy leukoplakia
59 (*7, 8*). The molecular pathogenesis in the nasopharyngeal epithelium is less clear as
60 experimental models of the human nasopharyngeal airway epithelium have not been
61 developed for EBV (*3*). Other than stratified keratinocytes, almost half of the
62 nasopharyngeal epithelium is composed of pseudostratified respiratory epithelium which
63 consists of a variety of cell types (ciliated, mucosecretory, basal and suprabasal) (*9*). In
64 this study, we present a *de novo* EBV infection model of the nasopharyngeal
65 pseudostratified epithelium grown in 3-D. Nasal primary cell cultures are differentiated
66 from conditionally reprogrammed cells of the human nasopharynx and grown in 3-D air-
67 liquid interface (ALI) culture (*10-12*). To distinguish this type of pseudostratified ALI
68 culture from other types of ALI culture with cell lines or organotypic rafts, we herein refer
69 to the primary nasal cell pseudostratified ALI as “pseudo-ALI” culture. Conventionally,

70 pseudo-ALI cultures of airway epithelial cells are used to study acute virus infections
71 such as influenza virus (*13*), respiratory syncytial virus (*14*), other respiratory pathogens
72 (*15*), and more recently SARS-CoV-2 (*16*). Here, we report that the pseudo-ALI culture
73 can be applied to the study of a persistent γ-herpesviruses, EBV. Using primary cells
74 from a collection of 9 donors, examples of both latent and lytic infection are observed.
75 Evidence of variation in donor susceptibility is presented. We report on one of the first
76 examples of EBV latent infection captured in primary nasopharyngeal cells. These
77 latently-infected cells are only observed in select donors, suggesting that some
78 individuals could harbour a latent, local reservoir.
79

80 **RESULTS**

81 **Establishment of a 3-D pseudo-ALI model of *de novo* EBV infection.** We have
82 previously demonstrated that conventional ALI culture can reactivate EBV from the NPC
83 cell line, HK1-EBV, to yield high infectious titres (>10^6 infectious green Raji units per
84 cm^2) (*17*). To elucidate EBV pathogenesis in primary cells, a method was developed for
85 *de novo* EBV infection in differentiated nasal epithelial cells in pseudo-ALI culture.
86 Primary cells from the nasopharynx, at the site of the lymphoid-rich Fossa of
87 Rosenmüller, were collected under direct visualization from adult immune-competent
88 donors undergoing endoscopic nasal procedures for reasons other than cancer. Primary
89 cells were expanded on irradiated mouse 3T3-J2 fibroblasts in the presence of ROCK
90 inhibitor (Y-27632) and lifted to the air-liquid interface on collagen-coated transwell
91 membranes for 4 weeks (*10, 18*). Once primary cells have differentiated into pseudo-
92 ALI cultures, EBV inoculum was applied to the apical surface by co-culture with the

93 EBV-positive Akata cell line reactivated with anti-human IgG. The producer Akata cell
94 line is recombinantly-infected with EBV expressing neomycin resistance and the EGFP
95 marker gene inserted into the non-essential *BXLF1*, herein referred to as rAkata (*19*).
96 As mock control, target cells were co-cultured with EBV-negative Akata cells similarly
97 treated with anti-human IgG antibody.
98
99 Cells differentiated in pseudo-ALI culture were analyzed by histopathology to control for
100 differentiation into respiratory epithelium (Supplementary Fig. 1). Hematoxylin and eosin
101 stain demonstrated the presence of pseudostratified epithelium and ciliated cells. Alcian
102 blue and periodic acid Schiff stain revealed mucin-secreting cells.
103 Immunohistochemistry staining for the proliferation marker, Ki67, showed the infrequent
104 presence of cycling cells in the basal layer. To identify susceptible samples, EBV
105 molecular diagnostics for latent and lytic markers of infection were developed for whole-
106 mount staining of pseudo-ALI culture. These molecular diagnostics were first validated
107 in the HK1-EBV cell line, in which 2-D culture is strictly latent but 3-D ALI culture
108 triggers lytic reactivation. The detection of EBV-encoded RNAs (EBERs) by *in situ*
109 hybridization (EBER-ISH) in the nucleus identifies latently-infected cells, while
110 immunofluorescence staining for Zebra (immediate-early protein) in the nucleus or
111 gp350 (late glycoprotein) in the cytoplasm denotes lytic infection (Supplementary Fig.
112 2). Notably, EBERs are not detected by EBER-ISH in oral hairy leukoplakia, a
113 permissive EBV infection, and are thus a diagnostic marker of latent infection (*20*).
114 Staining for the EBV oncoprotein, LMP1, identifies both latent and lytic infection.
115 Stained images are scored by pixel intensity represented as a histogram compared to

116 the mock (Fig. 1). Punctate LMP1 foci can also be discriminated as particles and scored
117 for particle intensity, represented as a box and whisker plot (Fig. 1C).
118
119 **EBV infection in pseudo-ALI culture show variation in donor susceptibility.** Both
120 susceptible and non-susceptible cultures were identified by EBV molecular diagnostics
121 (Table 1). A total of 3 pseudo-ALI cultures (donor no. 1, 4, 7) were susceptible to EBV
122 infection, while cultures from the other 6 donors were negative for the tested EBV
123 molecular markers (Fig. 1A, Supplementary Fig. 3). Pseudo-ALI cultures from 2 donors
124 (nos. 1 and 7) were positive for markers of latent infection, while cultures from donor no.
125 4 were positive for markers of lytic infection (Fig. 1A-C, Table 1). Stitched images
126 showed no evidence of residual B-cell contamination from the inoculum after
127 immunostain processing (Supplementary Fig. 4). In some cases, susceptible and non-
128 susceptible cultures could be identified in the same experiment using the same stock of
129 reactivated inoculum (Table 1). Thus, a failure to infect was indicative of host variation.
130 Infections were repeated on low-passaged cells thawed from banked conditionally
131 reprogrammed cells. In almost all cases of biological repeats (53 out of 54), either from
132 susceptible (donor no. 4 and 7) or non-susceptible (donor no. 3, 5 and 8) donors, the
133 same result in susceptibility and latent/lytic profiles were observed (Table 1,
134 parentheses). Susceptibility to EBV did not appear to correlate with the presence or
135 absence of comorbidity, although the number of samples collected is too small for
136 statistical analysis. The EBV entry receptor for epithelial cells, ephrin receptor A2
137 (EphA2) (*21, 22*), was detected on the plasma membrane in all susceptible pseudo-ALI
138 cultures but some of the cultures from non-susceptible samples were also strongly

139 positive for EphA2 staining, such as donors no. 3 and 6 (Supplementary Fig. 1, Table
140 1). This indicates that while expression of EphA2 is consistent with EBV infection, other
141 host factors also dictate susceptibility.
142
143 **Molecular diagnosis of EBV infection reveals donor-specific differences in**
144 **molecular pathogenesis – latent *vs.* lytic infection.** Samples from donors no. 4 and 7
145 were subjected to more extensive analyses at days 2 and 5 post-infection (p.i.). Donor
146 sample no. 4 stained positive for Zebra and LMP1 beginning at day 2 p.i., followed by
147 gp350 at day 5 p.i., denoting a lytic infection (Fig. 1B). Donor sample no. 7 showed
148 positivity for EBERs at day 5 p.i., denoting a latent infection (Fig. 1C). For donor sample
149 no. 4, EBV replication was measured by qPCR of DNA harvested from extracellular or
150 cell-associated DNase-resistant encapsidated virus (Supplementary Tables, A). As
151 input control, pseudo-ALI cultures were fixed before co-culture with the inoculum. While
152 the EBV genome copy number in the input control did not increase from day 2 to 5 p.i.,
153 extracellular EBV increased 37-fold (3.13×10^4 copies at day 2 p.i. to 1.16×10^6 copies
154 at day 5 p.i.). EBV copy numbers did not increase in the cell-associated virus which
155 measured between 1.55 – 4.06×10^4 copies. This indicates that the majority of
156 encapsidated EBV are packaged for secretion. Using virus collected from the
157 extracellular source, infectious units were scored by the Green Raji Unit (GRU) assay in
158 the non-producer Raji cell line. The secreted virus is indeed infectious, reaching $1.07 \times$
159 10^5 GRUs by 5 days p.i. (Supplementary Tables, A).
160

161 **Single cell RNA-sequencing (scRNA-seq) reveals cell type-specific EBV**
162 **transcriptional profiles.** scRNA-seq analysis poses a challenge for all herpesvirus
163 genomes because of overlapping 3' co-terminal herpesvirus transcripts, whose non-
164 uniquely mapped reads are discarded in the 10X Genomics single cell analysis pipeline
165 (*23*). We reasoned however that this bioinformatics challenge is theoretically possible
166 with the EBV gamma-herpesvirus genomes given that it has been demonstrated for
167 alpha- and beta-herpesviruses (*24-26*). To identify EBV-infected cell types, donor
168 sample no. 4 was subjected to scRNA-seq. Cell clusters (Fig. 2A), were assigned cell
169 identities using experimentally-defined marker genes defined from primary human nasal
170 epithelial cells grown in pseudo-ALI culture (*27*) and primary nasal tissue (The Human
171 Cell Atlas Lung Consortium) (*28*), (Supplementary Fig. 5). All major airway epithelial cell
172 types (basal, secretory, suprabasal and ciliated) could be identified (Fig. 2A). In order to
173 improve alignment to the partially annotated EBV genome (NCBI KC207813.1), the
174 reference genome for the Akata strain was updated with additional exon annotation
175 totaling 87 genes. As there are no reports of scRNA-seq analysis on the EBV
176 transcriptome, we tested several alignment algorithms using the 10X Genomics Cell
177 Ranger pipeline. The reads were either aligned to the whole EBV genome as one
178 annotation, as separately annotated genes, or as annotated genes but with genes that
179 have regions of overlap in the same direction represented as fusion genes. Alignment to
180 the separate annotation assigns the identity of EBV transcripts according to the
181 reference annotations, but alignment to the other two annotations counts more EBV
182 reads. Overall, the EBV transcriptome represents 0.08% (separate annotation) to 0.17%
183 (one annotation and fused annotation) of the total transcriptome (Fig. 2B). This is similar

184 to estimates from bulk RNA-seq of lymphoblastoid cell lines carrying latent EBV, where
185 the majority of samples had EBV reads measuring 0.1-0.5% of the total transcriptome
186 (*29*). A large majority of the cells (71%-82%) expressed EBV and/or EGFP transcripts
187 (Fig. 2B).
188
189 Every cluster scored positive for EBV and/or EGFP reads (Fig. 2C). *BHLF1, BHRF1,*
190 *LF3 and LMP1/BNLF2a/BNLF2b* were the most frequently detected genes in the
191 highest proportion of epithelial cells across clusters (Supplementary Fig. 6). All the cells
192 in cluster 4 defined as the B-cell inoculum, expressed EBV and/or EGFP transcripts,
193 with >97% of cells showing both EBV and EGFP (Fig. 2C). Across the epithelial cell
194 clusters (clusters 1,2,3,5,6,7,0) the percent of cells with EBV reads ranged between
195 63%-91%, with no clear difference in susceptibility between clusters (Fig. 2C). However,
196 density plots revealed two distinct EBV expression profiles, clusters with a peak at low
197 UMI count ($\log_{10}$(count+1) < 0.3, clusters 0,1,3,5,7) denoted as EBV^{low}, and clusters
198 with 1-3 $\log_{10}$ higher UMI counts (clusters 2,4,6) denoted as EBV^{high} (Fig. 2D).
199
200 **Lytic infection is localized to suprabasal cells while latent infection is confined to**
201 **basal/secretory and ciliated cell types.** EBV^{low} cells found in all clusters displayed a
202 distinct expression pattern (*BHLF1, BHRF1, LF3,* and the fused annotation *LMP-*
203 *1/BNLFa/BNLFb*) which did not resemble a canonical type I/II/III latency profile
204 (Supplementary Fig. 7). These cells are likely to be latent, refractory or in the early
205 stage of the lytic cascade. These EBV^{low} cells are predominantly found in basal,
206 secretory and ciliated cell types but also in a group of suprabasal cells defined by

207 cluster 0 (Supplementary Fig.7). The mixed cell types in Cluster 6 could be further
208 divided into 4 sub-populations with distinct marker gene expression (Supplementary Fig.
209 8A). EBV^{high} cells within sub-cluster 6-4 with basal cell features (Supplementary Fig. 8)
210 and the EBV^{high} suprabasal cells in cluster 0 (Supplementary Fig. 7) display high levels
211 of *BZLF1/BRLF1* indicative of reactivation. Lytic infection is mainly observed in EBV^{high}
212 suprabasal cells (cluster 2) where there is global induction of EBV genes
213 (Supplementary Fig. 7) but shut-off of host mRNA (Supplementary Fig. 9). Overall, all
214 cell types in the pseudo-ALI were susceptible to EBV-infection; however productive
215 virus infection is mainly confined to suprabasal cells. Despite the concerns of
216 overlapping and low abundance viral transcripts evading capture by scRNA-seq (*23*),
217 we conclude that with the appropriate reference annotation, the γ-herpesvirus genome
218 can be analyzed by scRNA-seq.
219

220 **DISCUSSION**

221 In conclusion, we demonstrate that pseudostratified epithelial cells from the
222 nasopharynx are susceptible to EBV infection. Results from this study would indicate
223 that host variables other than the expression of EphA2 impact susceptibility in the
224 nasopharynx as well as the type of EBV infection (productive or non-productive). Given
225 the relatively rare chance of finding an EBV-infected nasopharyngeal cell in
226 asymptomatic carriers (*5, 6*), the pseudo-ALI culture thus provides a new organoid
227 model in which to study EBV molecular pathogenesis in the nasopharynx. Our findings
228 agree with prior studies in organotypic rafts using oral-derived keratinocytes that EBV
229 lytic infection is confined to suprabasal cells (*8, 30*). We conclude that latent infection

230 can occur in nasopharynx-derived basal/secretory/ciliated cell types, which may habour
231 a local EBV reservoir, and suggest that the basal cell type could be a progenitor cell for
232 NPC tumors.
233

234 **MATERIALS AND METHODS**

235 **Samples.** Primary nasopharyngeal cell samples were collected at UPMC Mercy
236 hospital before emergence of the COVID-19 pandemic conducted under IRB
237 STUDY#19030014 University of Pittsburgh Sinus Fluid and Tissue Bank. Voluntary
238 informed consent was obtained for the collection, storage and analysis of biologic
239 and/or genetic material for research, and such de-identified samples and de-identified
240 data may be shared with other investigators for health research.
241

242 **Cell culture.** The HK1 NPC cell line and the Akata Burkitt's lymphoma B-cell line were
243 maintained in RPMI supplemented with 10% fetal bovine serum. HK1 and Akata cells
244 infected with the EBV recombinant Akata strain (courtesy of Dr. George Tsao, Hong
245 Kong University) were supplemented with 800 μg/mL G418 selection(*19, 31*). The EBV-
246 infected HK1 (HK1-EBV) and Akata (rAkata) cells express neomycin-resistance and
247 EGFP from the SV40 early promoter, inserted into the EBV *BXLF1* locus, and are intact
248 for expression of the EBV miRNAs(*19, 32*). Cells were incubated at 37°C with 5% CO_2
249 and confirmed to be negative for mycoplasma contamination by PCR. Primary nasal
250 epithelial cells were cultured from nasal cytobrush scrapings of the nasopharynx or
251 inferior turbinate. Protocols for Collected cells were seeded on irradiated mouse 3T3-J2
252 feeder fibroblasts and expanded in Georgetown media(*10*). The presence of 4 μM

253 ROCK inhibitor (Y-27632) extends the lifespan and induces the conditional
254 reprogramming of epithelial cells(*18*). Media was changed daily, and cells were sub-
255 cultured at 1:4 seeding density. At passage 1 or 2, $1.5x10^5$ cells were seeded on human
256 type IV placental collagen-coated transwell filters (Corning, 0.33 cm^2, 0.4 µm,
257 polyethylene terephthalate) in Georgetown media for 24 hours. After 24 hours apical
258 media was removed, cultures washed once in PBS, and the basolateral media was
259 replaced with 400 µL of ALI medium(*11*) supplemented with 0.5% Ultroser G Serum
260 Substitute (PALL), denoted as UNC/USG basolateral media. Cultures were
261 maintained at the air-liquid interface for at least 4 weeks to allow differentiation into a
262 pseudo-ALI culture. Basolateral media was changed 3 times a week. HK1 and HK1-
263 EBV cells were cultured at the air-liquid interface as previously described(*17*).
264
265 **EBV Infection.** rAkata EBV-infected cells was reactivated at $1x10^6$ cells/mL with a goat
266 polyclonal anti-human IgG Fc-specific antibody (Sigma) for 48 hours. EBV-negative
267 Akata cells were similarly treated with anti-human IgG antibody as a mock control. Virus
268 production was confirmed by quantitative PCR for *BALF5*, as described in
269 Supplementary Methods. Reactivated Akata cells were pelleted by centrifugation and
270 resuspended at a concentration of $1.25x10^7$ cells /mL in calcium-/magnesium-free
271 Dulbecco's PBS (DPBS). Primary pseudo-ALI cultures were washed in DPBS once for 5
272 minutes at 37°C and twice briefly at room temperature. The reactivated B-cell
273 suspension was added to the apical surface of the pseudo-ALI culture in 200 µL,
274 basolateral media was replaced with DPBS, and cultures were pre-incubated at 37°C
275 for 2 hours. The basolateral DPBS was then replaced with UNC/USG media and

276 cultures incubated for a further 48 hours at 37°C. B-cell co-culture was removed by
277 aspiration, and pseudo-ALI cultures were washed three times in Hank's buffered saline
278 solution (HBSS) to remove remaining B-cells. Cultures were fixed (2 days p.i.) or
279 incubated at 37°C for up to 5 additional days (4-7 days p.i.), changing UNC/USG
280 basolateral media every 48 hours.
281
282 **scRNA-seq.** Cell suspensions were loaded into 10X Genomics Chromium instrument
283 for library preparation as described previously (*33*), using the single cell 3'v3.1
284 (SC3Pv3) chemistry. Library QC was performed on an Agilent Bioanalyzer. High-
285 throughput sequencing was performed by Novogene on a HiSeq paired-end 150 bp
286 configuration yielding >472M reads.
287
288 **Code availability.** The R script for Seurat workflow and for data visualization is
289 available upon request.
290
291 **Data availability.** Raw and processed scRNA-seq data files, and the merged
292 EBV+hg38 genome annotation file will be deposited in NCBI Gene Expression Omnibus
293 (GEO). Filtering criteria and data processing steps are provided in the GEO submission.
294 The Akata EBV genome (NCBI KC207813.1) with updated annotations are available in
295 Github (https://github.com/TangLabGOT/Reference-Genomes).
296
297 **ACKNOWLEDGEMENTS**

298 This work is supported in part by the Hillman Foundation and used the Hillman Tissue
299 and Research Pathology Services shared resource that is supported in part by National
300 Institutes of Health award P30CA047904. We thank Tracy Tabib and Dr. Robert Lafyatis
301 from the University of Pittsburgh Single Cell Core for advice on scRNA-seq and Dr.
302 George S.W. Tsao (Hong Kong University, Hong Kong SAR) for providing the HK1 and
303 HK1-EBV cell lines. We thank the Cystic Fibrosis Research Center cell core at the
304 University of Pittsburgh (funded by the Cystic Fibrosis Foundation Research
305 Development Program to the University of Pittsburgh) for providing reagents and advice
306 on the culture of primary cells. We thank the Bioinformatics Core Facility at the
307 Sahlgrenska Academy for bioinformatics support.

308

309 **AUTHOR CONTRIBUTIONS**

310 K.S. designed experiments, analysed data and prepared the manuscript. P.Z.
311 conducted experiments for EBV infection and molecular diagnosis. S.L. and J.M.
312 collected specimens, comorbidity information and revised the manuscript. M.M.
313 provided protocols and reagents for pseudo-ALI culture and revised the manuscript.
314 P.Z., Y.B., Y.T., S.A., H.J.P. and K-W.T. analysed data and prepared the manuscript.
315 A.G. established protocols and performed the histology on pseudo-ALI cultures.

316

317 **COMPETING INTERESTS**

318 The authors declare no competing interests.

319

320 **FIGURE LEGENDS**

321 Figure 1. EBV *de novo* infection of primary nasopharynx-derived epithelial cells in
322 pseudo-ALI culture. A-C, Immunofluorescence staining and EBER-ISH images (red) for
323 EBV molecular diagnostics in pseudo-ALI cultures infected with EBV-positive rAkata B-
324 cells or EBV-negative Akata B-cells (mock), counterstained with DAPI (blue). Shown
325 are maximum intensity projections of confocal images on the *xy* (square) and *xz*
326 (rectangle) planes. The pixel signal intensity of the EBV markers in the unzoomed
327 image, in mock (blue line) and EBV-infected (red line) samples, are compared in the
328 histograms. (A) Shown are the results for four donors. Cells were harvested at days 4-5
329 p.i. for Zebra and days 5-7 p.i. for EBER-ISH. More extensive analysis at days 2 and 5
330 p.i. was performed for donors (B) no. 4 and (C) no. 7. LMP1 foci in the weakly stained
331 sample (donor no. 7) was also scored as mean particle intensity, from the unzoomed
332 image. Bold, susceptible donor sample; italics, non-susceptible donor sample. LUTs,
333 look-up-tables.

334

335 Figure 2. Identification of EBV-infected cell types from scRNA-seq performed on a
336 pseudo-ALI culture from donor no. 4. (A) Uniform Manifold Approximation and
337 Projection (UMAP) plots displaying the major cell clusters and assigned identity from
338 host marker genes. Schematic displays the cell types in the pseudostratified nasal
339 epithelium that are represented in the cell clusters. (B) Comparison of alignment
340 methods against different EBV genome annotations by UMI counts and numbers of
341 cells. (C) The percentage of EBV-infected cells is displayed for each cluster. Shown are
342 the results from alignment to the fused annotation EBV genome, (D) Density plots of

343 log10 transformed pseudocount (UMI count+1) for EBV reads against normalized cell
344 numbers displayed for each cell cluster. Cell numbers are normalized for the total
345 number of cells in each cluster. Cells are grouped by $EBV^{negative}$, EBV^{low} and EBV^{high}
346 populations. Density plots with a similar profile are similarly coloured.
347
348 TABLE 1. Summary of molecular diagnostic results from EBV infection of nasal pseudo-
349 ALIs with rAkata.
350

351 **REFERENCES**

352 1. L. S. Young, L. F. Yap, P. G. Murray, Epstein-Barr virus: more than 50 years old and still
353 providing surprises. *Nature reviews. Cancer* **16**, 789-802 (2016).
354 2. N. Raab-Traub, Nasopharyngeal Carcinoma: An Evolving Role for the Epstein-Barr Virus.
355 *Current topics in microbiology and immunology* **390**, 339-363 (2015).
356 3. K. H. Y. Shair, A. Reddy, V. S. Cooper, in *Cancers (Basel)*. (2018), vol. 10.
357 4. S. W. Tsao *et al.*, The biology of EBV infection in human epithelial cells. *Seminars in cancer*
358 *biology* **22**, 137-143 (2012).
359 5. R. Pathmanathan, U. Prasad, R. Sadler, K. Flynn, N. Raab-Traub, Clonal proliferations of
360 cells infected with Epstein-Barr virus in preinvasive lesions related to nasopharyngeal
361 carcinoma. *The New England journal of medicine* **333**, 693-698 (1995).
362 6. C. K. Sam *et al.*, Analysis of Epstein-Barr virus infection in nasopharyngeal biopsies from a
363 group at high risk of nasopharyngeal carcinoma. *International journal of cancer. Journal*
364 *international du cancer* **53**, 957-962 (1993).
365 7. L. M. Hutt-Fletcher, The Long and Complicated Relationship between Epstein-Barr Virus
366 and Epithelial Cells. *Journal of virology* **91**, (2017).
367 8. R. M. Temple *et al.*, Efficient replication of Epstein-Barr virus in stratified epithelium in
368 vitro. *Proceedings of the National Academy of Sciences of the United States of America*,
369 (2014).
370 9. M. Y. Ali, Histology of the human nasopharyngeal mucosa. *J Anat* **99**, 657-672 (1965).
371 10. F. Serrano Castillo *et al.*, A physiologically-motivated model of cystic fibrosis liquid and
372 solute transport dynamics across primary human nasal epithelia. *J Pharmacokinet*
373 *Pharmacodyn* **46**, 457-472 (2019).
374 11. M. L. Fulcher, S. H. Randell, in *Methods in molecular biology*. (2013), vol. 945, pp. 109-
375 121.
376 12. S. Chen, J. Schoen, Air-liquid interface cell culture: From airway epithelium to the female
377 reproductive tract. *Reprod Domest Anim* **54 Suppl 3**, 38-45 (2019).

378 13. M. Richard *et al.*, Influenza A viruses are transmitted via the air from the nasal respiratory
379 epithelium of ferrets. *Nature communications* **11**, 766 (2020).
380 14. C. S. Anderson *et al.*, CX3CR1 as a respiratory syncytial virus receptor in pediatric human
381 lung. *Pediatr Res* **87**, 862-867 (2020).
382 15. W. Hao *et al.*, Infection and propagation of human rhinovirus C in human airway epithelial
383 cells. *Journal of virology* **86**, 13524-13532 (2012).
384 16. R. Lu *et al.*, Genomic characterisation and epidemiology of 2019 novel coronavirus:
385 implications for virus origins and receptor binding. *Lancet* **395**, 565-574 (2020).
386 17. E. A. Caves *et al.*, Air-Liquid Interface Method To Study Epstein-Barr Virus Pathogenesis in
387 Nasopharyngeal Epithelial Cells. *mSphere* **3**, (2018).
388 18. X. Liu *et al.*, ROCK inhibitor and feeder cells induce the conditional reprogramming of
389 epithelial cells. *The American journal of pathology* **180**, 599-607 (2012).
390 19. S. Maruo, L. Yang, K. Takada, Roles of Epstein-Barr virus glycoproteins gp350 and gp25 in
391 the infection of human epithelial cells. *The Journal of general virology* **82**, 2373-2383
392 (2001).
393 20. K. Gilligan, P. Rajadurai, L. Resnick, N. Raab-Traub, Epstein-Barr virus small nuclear RNAs
394 are not expressed in permissively infected cells in AIDS-associated leukoplakia.
395 *Proceedings of the National Academy of Sciences of the United States of America* **87**, 8790-
396 8794 (1990).
397 21. J. Chen *et al.*, Ephrin receptor A2 is a functional entry receptor for Epstein-Barr virus. *Nat*
398 *Microbiol* **3**, 172-180 (2018).
399 22. H. Zhang *et al.*, Ephrin receptor A2 is an epithelial cell receptor for Epstein-Barr virus
400 entry. *Nat Microbiol* **3**, 164-171 (2018).
401 23. D. P. Depledge, I. Mohr, A. C. Wilson, Going the Distance: Optimizing RNA-Seq Strategies
402 for Transcriptomic Analysis of Complex Viral Genomes. *Journal of virology* **93**, (2019).
403 24. N. Drayman, P. Patel, L. Vistain, S. Tay, HSV-1 single-cell analysis reveals the activation of
404 anti-viral and developmental programs in distinct sub-populations. *Elife* **8**, (2019).
405 25. E. Wyler *et al.*, Single-cell RNA-sequencing of herpes simplex virus 1-infected cells
406 connects NRF2 activation to an antiviral program. *Nature communications* **10**, 4878
407 (2019).
408 26. M. Shnayder *et al.*, Single cell analysis reveals human cytomegalovirus drives latently
409 infected cells towards an anergic-like monocyte state. *Elife* **9**, (2020).
410 27. S. Ruiz Garcia *et al.*, Novel dynamics of human mucociliary differentiation revealed by
411 single-cell RNA sequencing of nasal epithelial cultures. *Development* **146**, (2019).
412 28. F. A. Vieira Braga *et al.*, A cellular census of human lungs identifies novel cell states in
413 health and in asthma. *Nature medicine* **25**, 1153-1163 (2019).
414 29. A. Arvey *et al.*, An atlas of the Epstein-Barr virus transcriptome and epigenome reveals
415 host-virus regulatory interactions. *Cell host & microbe* **12**, 233-245 (2012).
416 30. D. M. Nawandar *et al.*, Differentiation-Dependent KLF4 Expression Promotes Lytic
417 Epstein-Barr Virus Infection in Epithelial Cells. *PLoS pathogens* **11**, e1005195 (2015).
418 31. A. K. Lo *et al.*, Epstein-Barr virus infection alters cellular signal cascades in human
419 nasopharyngeal epithelial cells. *Neoplasia* **8**, 173-180 (2006).

420 32. A. K. Lo *et al.*, Modulation of LMP1 protein expression by EBV-encoded microRNAs.
421 *Proceedings of the National Academy of Sciences of the United States of America* **104**,
422 16164-16169 (2007).
423 33. C. Morse *et al.*, Proliferating SPP1/MERTK-expressing macrophages in idiopathic
424 pulmonary fibrosis. *Eur Respir J* **54**, (2019).
425

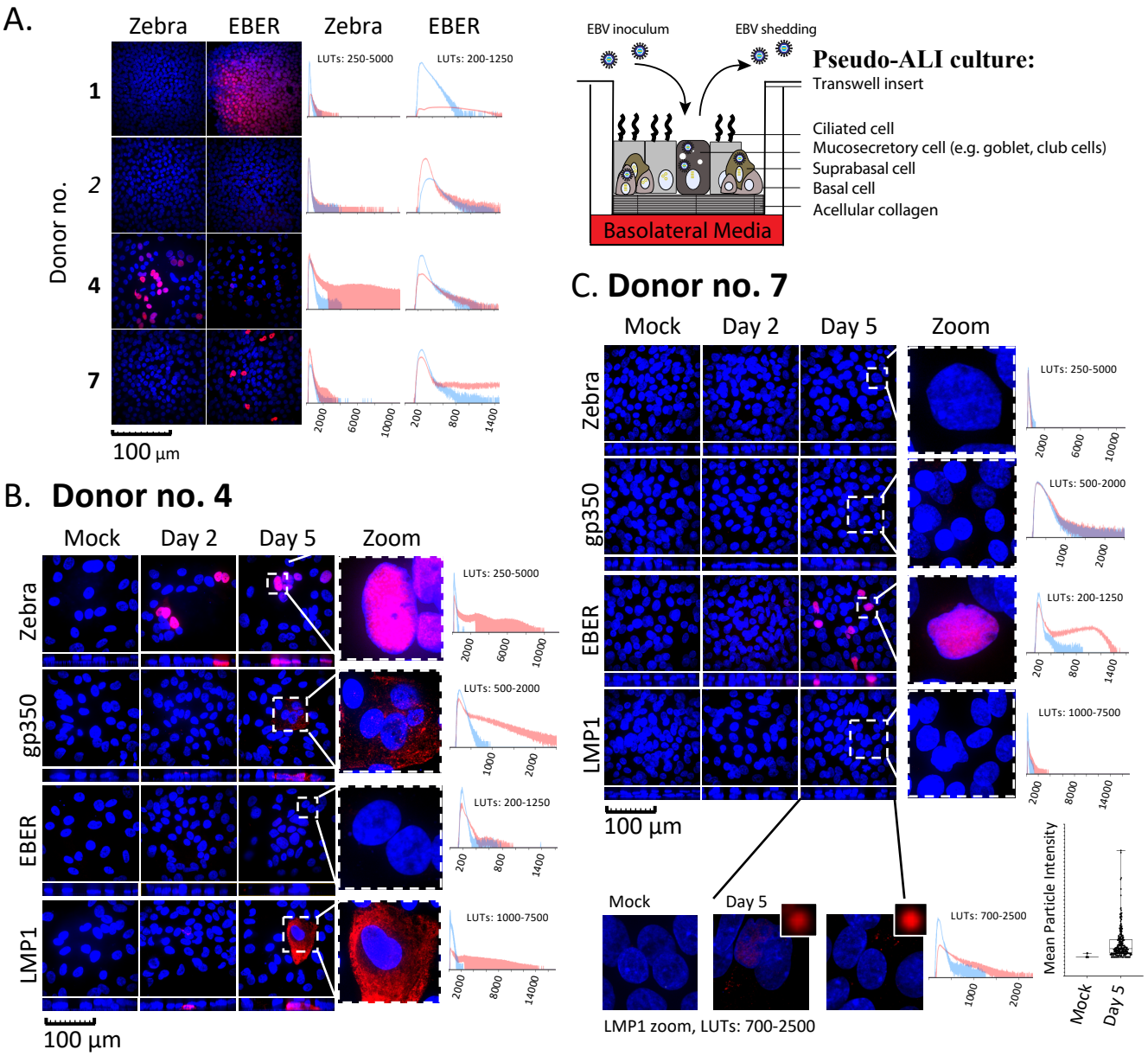
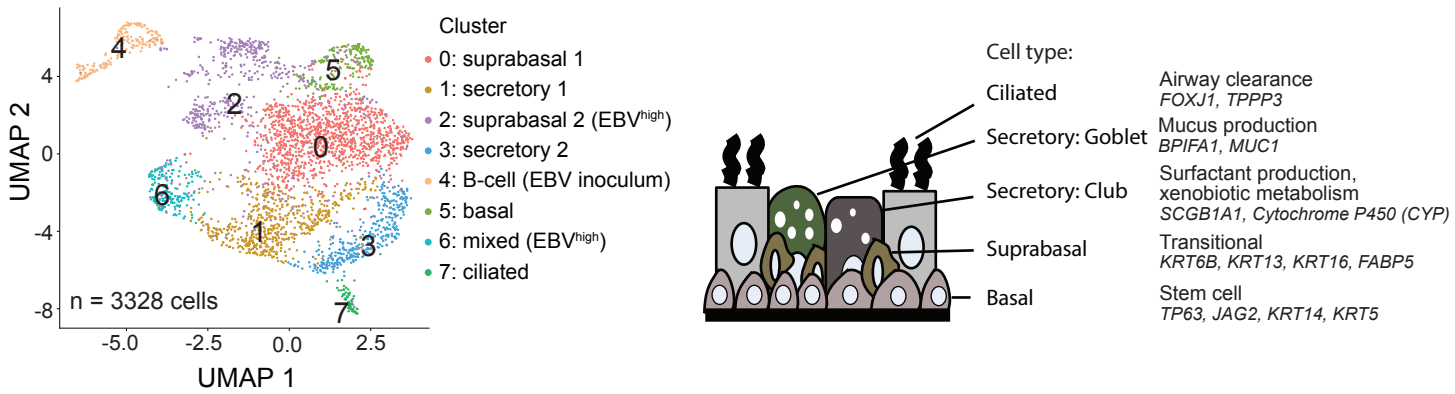
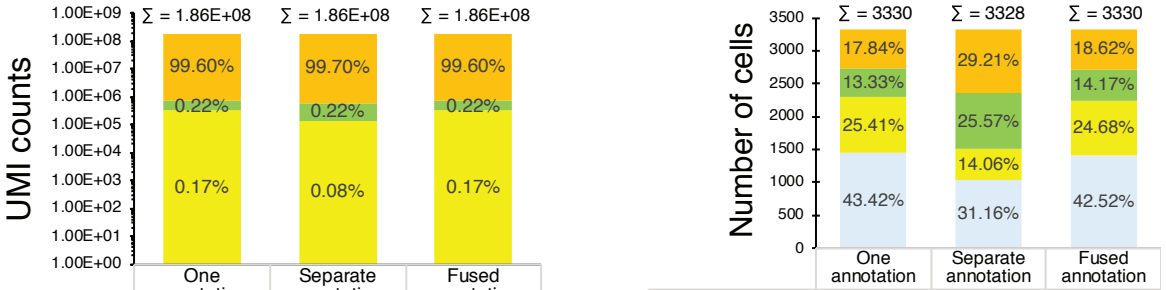


Figure 1. EBV *de novo* infection of primary nasopharynx-derived epithelial cells in pseudo-ALI culture. A-C, Immunofluorescence staining and EBER-ISH images (red) for EBV molecular diagnostics in pseudo-ALI cultures infected with EBV-positive rAkata B-cells or EBV-negative Akata B-cells (mock), counterstained with DAPI (blue). Shown are maximum intensity projections of confocal images on the *xy* (square) and *xz* (rectangle) planes. The pixel signal intensity of the EBV markers in the unzoomed image, in mock (blue line) and EBV-infected (red line) samples, are compared in the histograms. **(A)** Shown are the results for four donors. Cells were harvested at days 4-5 p.i. for Zebra and days 5-7 p.i. for EBER-ISH. More extensive analysis at days 2 and 5 p.i. was performed for donors **(B)** no. 4 and **(C)** no. 7. LMP1 foci in the weakly stained sample (donor no. 7) was also scored as mean particle intensity, from the unzoomed image. Bold, susceptible donor sample; italics, non-susceptible donor sample. LUTs, look-up-tables.

A



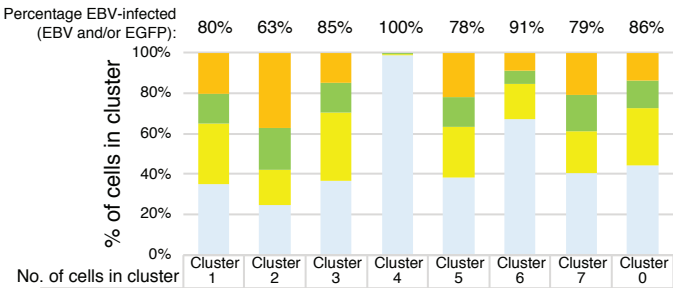
B



	One annotation	Separate annotation	Fused annotation
Uninfected	1.86E+08	1.86E+08	1.85E+08
EGFP	4.18E+05	4.18E+05	4.18E+05
EBV	3.25E+05	1.41E+05	3.21E+05

	One annotation	Separate annotation	Fused annotation
Uninfected	594	972	620
EGFP+ only	444	851	472
EBV+ only	846	468	822
EBV+ and EGFP+	1446	1037	1416

C



No. of cells in cluster	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 0
Uninfected	120	211	54	0	41	16	12	164
EGFP+ only	90	117	55	1	28	12	10	161
EBV+ only	180	96	124	1	46	32	12	331
EBV+ and EGFP+	208	140	135	190	72	123	23	523

D

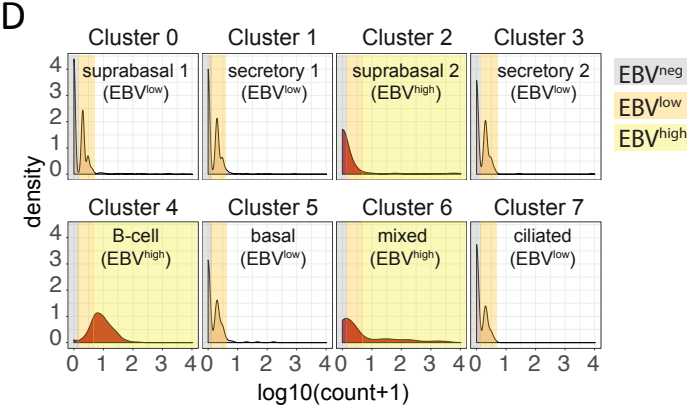


Figure 2. Identification of EBV-infected cell types from scRNA-seq performed on a pseudo-ALI culture from donor no. 4. (**A**) Uniform Manifold Approximation and Projection (UMAP) plots displaying the major cell clusters and assigned identity from host marker genes. Schematic displays the cell types in the pseudostratified nasal epithelium that are represented in the cell clusters. (**B**) Comparison of alignment methods against different EBV genome annotations, by UMI counts and numbers of cells. (**C**) The percentage of EBV-infected cells is displayed for each cluster. Shown are the results from alignment to the fused annotation EBV genome, (**D**) Density plots of log10 transformed pseudocount (UMI count+1) for EBV reads against normalized cell numbers displayed for each cell cluster. Cell numbers are normalized for the total number of cells in each cluster. Cells are grouped by EBV^{negative}, EBV^{low} and EBV^{high} populations. Density plots with a similar profile are similarly coloured.

Table 1. Summary of molecular diagnostic results from EBV infection of nasal pseudo-ALIs with rAkata.

Donor No.* (reference code)	Experiment No.	Susceptibility to EBV (-, +)	EphA2**	Latent	Lytic (immediate-early, late)		Latent-lytic	Reason for surgery/ comorbidity [other]
				EBER (-/+)	Zebra (-/+)	gp350 (-/+)	LMP1 (-/+)	
1 (H04)	1	+	++	+	-	n/a	n/a	CRSsNP***/none
2 (H05)		-	+	-	-	n/a	n/a	Fungus ball, septal deviation, allergic rhinitis/inflammatory polyarthropathy
3 (H08)	2	-	+++	- (6/6)	- (6/6)	n/a	n/a	CRSsNP/inflammatory polyarthropathy, Yellow Nail Syndrome
4 (H10)	3	+	++	-	+ (5/6)	+	+	Left turbinate hypertrophy/none
5 (H12)	4	-	+	- (6/6)	- (6/6)	-	-	CRSsNP/inflammatory polyarthropathy
6 (H13)		-	+++	-	-	-	-	Silent sinus syndrome/thyroid disorder-unspecified, [Crohn's disease]
7 (H15)	5	+	++	+ (6/6)	- (6/6)	-	+, weak	Septal perforation-unknown etiology/[rheumatoid arthritis]
8 (H17)	6	-	-	- (6/6)	- (6/6)	-	-	Septal deviation/none
9 (H18)		-	+	-	-	-	-	Odontogenic sinusitis and septal deviation/none

*Donor cells were obtained from the nasopharynx. The reference code denotes a unique donor identifier; **Scoring criteria: -.+,++,+++; ***CRSsNP, Chronic rhinosinusitis without nasal polyps; n/a, not available; biological repeats are indicated in parentheses.