Ectopic expression of HIV-1 Tat modifies gene expression in cultured B cells: Implications for the development of B-cell lymphomas in HIV-1-infected patients

Anna A. Valyaeva [1, 2, 3] , **Maria A. Tikhomirova** [1, 4] , **Daria M. Potashnikova** [1] , **Alexandra N. Bogomazova** [5, 6] , **Galina P. Snigiryova** [5] , **Aleksey A. Penin** [7] , **Maria D. Logacheva** [1, 8] , **Eugene A. Arifulin** [1] , **Anna A. Shmakova** [4, 9] , **Diego Germini** [9] , **Anastasia I. Kachalova** [1] , **Aleena A. Saidova** [1, 10] , **Anastasia A. Zharikova** [1, 2] , **Yana R. Musinova** [1, 4] , **Andrey A. Mironov** [1, 7] , **Yegor S Vassetzky** [Corresp., 4, 9] , **Eugene V. Sheval** [Corresp. 1, 2]

[1] Lomonosov Moscow State University, Moscow, Russia

[2] Belozersky Institute of Physico-Chemical Biology, , Lomonosov Moscow State University, Moscow, Russia

[3] Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, Moscow, Russia

[4] Koltzov Institute of Developmental Biology, Moscow, Russia

[5] Federal Research and Clinical Center of Physical-Chemical Medicine, Moscow, Russia

[6] Center for Precision Genome Editing and Genetic Technologies for Biomedicine, ederal Research and Clinical Center of Physical-Chemical Medicine of Federal Medical Biological Agency, Moscow, Russia

[7] Institute for Information Transmission Problems, Moscow, Russia

[8] Skolkovo Institute of Science and Technology, Moscow, Russia

[9] UMR9018 (CNRS - Institut Gustave Roussy - Université Paris Saclay), Centre National de Recherche Scientifique, Villejuif, France, France

[10] 1Center for Precision Genome Editing and Genetic Technologies for Biomedicine, Engelhardt Institute of Molecular Biology, Moscow, Russia

Corresponding Authors: Yegor S Vassetzky, Eugene V. Sheval
Email address: yegor.vassetzky@cnrs.fr, sheval_e@belozersky.msu.ru

An increased frequency of B-cell lymphomas is observed in human immunodeficiency virus-1 (HIV-1)-infected patients, although HIV-1 does not infect B cells. Development of B-cell lymphomas may be potentially due to the action of the HIV-1 Tat protein, which is actively released from HIV-1-infected cells, on uninfected B cells. The exact mechanism of Tat-induced B-cell lymphomagenesis has not yet been precisely identified. Here, we ectopically expressed either Tat or its TatC22G mutant devoid of transactivation activity in the RPMI 8866 lymphoblastoid B cell line and performed a genome-wide analysis of host gene expression. Stable expression of both Tat and TatC22G led to substantial modifications of the host transcriptome, including pronounced changes in antiviral response and cell cycle pathways. We did not find any strong action of Tat on cell proliferation, but during prolonged culturing, Tat-expressing cells were displaced by non-expressing cells, indicating that Tat expression slightly inhibited cell growth. We also found an increased frequency of chromosome aberrations in cells expressing Tat. Thus, Tat can modify gene expression in cultured B cells, leading to subtle modifications in cellular growth and chromosome instability, which could promote lymphomagenesis over time.

1 Ectopic expression of HIV-1 Tat modifies gene expression in cultured B cells:

2 Implications for the development of B-cell lymphomas in HIV-1-infected patients

4 Anna A. Valyaeva [1,2,3], Maria A. Tikhomirova [1,4], Daria M. Potashnikova [3], Alexandra N.

5 Bogomazova [5,6], Galina P. Snigiryova [7], Aleksey A. Penin [8], Maria D. Logacheva [2,9], Eugene A.

6 Arifulin [2], Anna Shmakova [4,10], Diego Germini [10], Anastasia I. Kachalova [3], Aleena A. Saidova

7 [3,11], Anastasia A. Zharikova [1,2], Yana R. Musinova [2,4], Andrey A. Mironov [1,12], Yegor S.

8 Vassetzky [4,10] and Eugene V. Sheval [2,3*]

10 [1]Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, Moscow,

11 Russia

12 [2]Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University,

13 Moscow, Russia

14 [3]Department of Cell Biology and Histology, Faculty of Biology, Lomonosov Moscow State

15 University, Moscow, Russia

16 [4]Koltzov Institute of Developmental Biology, Russian Academy of Sciences, Moscow, Russia

17 [5]Federal Research and Clinical Center of Physical-Chemical Medicine of Federal Medical

18 Biological Agency, Moscow, Russia

19 [6]Center for Precision Genome Editing and Genetic Technologies for Biomedicine, Federal

20 Research and Clinical Center of Physical-Chemical Medicine of Federal Medical Biological

21 Agency, Moscow, Russia

22 [7]Burdenko National Medical Research Center of Neurosurgery, Moscow, Russia

23 [8]Laboratory of Plant Genomics, Institute for Information Transmission Problems, Russian

24 Academy of Sciences, Moscow, Russia

25 [9]Center of Molecular and Cellular Biology, Skolkovo Institute of Science and Technology,

26 Moscow, Russia

27 [10]CNRS, UMR 9018, Université Paris-Saclay, Institut Gustave Roussy, Villejuif, France

28 [11]Center for Precision Genome Editing and Genetic Technologies for Biomedicine, Engelhardt

29 Institute of Molecular Biology, Russian Academy of Sciences, Russia

30 [12]Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow,

31 Russia

32   **ABSTRACT**

33   An increased frequency of B-cell lymphomas is observed in human immunodeficiency virus-1

34   (HIV-1)-infected patients, although HIV-1 does not infect B cells. Development of B-cell

35   lymphomas may be potentially due to the action of the HIV-1 Tat protein, which is actively

36   released from HIV-1-infected cells, on uninfected B cells. The exact mechanism of Tat-induced

37   B-cell lymphomagenesis has not yet been precisely identified. Here, we ectopically expressed

38   either Tat or its TatC22G mutant devoid of transactivation activity in the RPMI 8866

39   lymphoblastoid B cell line and performed a genome-wide analysis of host gene expression.

40   Stable expression of both Tat and TatC22G led to substantial modifications of the host

41   transcriptome, including pronounced changes in antiviral response and cell cycle pathways. We

42   did not find any strong action of Tat on cell proliferation, but during prolonged culturing, Tat-

43   expressing cells were displaced by non-expressing cells, indicating that Tat expression slightly

44   inhibited cell growth. We also found an increased frequency of chromosome aberrations in cells

45   expressing Tat. Thus, Tat can modify gene expression in cultured B cells, leading to subtle

46   modifications in cellular growth and chromosome instability, which could promote

47   lymphomagenesis over time.

48   **Keywords** HIV-1 Tat, B cells, Virus-cell interactions, Gene expression, RNA-seq

INTRODUCTION

49

50 In the second part of the XX century, human immunodeficiency virus-1 (HIV-1) has rapidly

51 spread throughout the world and caused high mortality due to its high evolution rate. HIV-1

52 preferentially infects CD4$^+$ T cells, macrophages, and microglial cells, leading to a damaged

53 immune system and the development of acquired immunodeficiency syndrome (AIDS).

54 Combined antiretroviral therapy (cART) stops the virus from making copies of itself in the body

55 and may slow down the development of AIDS (Autran et al., 1997; HIV-CAUSAL

56 Collaboration et al., 2010). However, even after the introduction of cART, individuals infected

57 with HIV-1 are at significantly higher risk of developing non-AIDS-related comorbidities,

58 including the development of neurocognitive disorders (Mateen et al., 2012; Marino et al., 2020),

59 cardiovascular diseases (Wang et al., 2015; Jiang et al., 2018), adipose tissue senescence

60 (Gorwood et al., 2020), and cancer (Shmakova, Germini & Vassetzky, 2020).

61   Despite the improved control of HIV-1 infection achieved by cART, B-cell lymphomas are

62 still elevated in HIV-1-infected patients and are the most frequent cause of death in these patients

63 (Noy, 2020; Shmakova, Germini & Vassetzky, 2020; Hübel, 2020). However, even more

64 surprising is the fact that HIV-1-infected patients have an increased incidence of specific B-cell

65 lymphomas, namely Burkitt lymphoma and diffuse large B-cell lymphoma (Gloghini, Dolcetti &

66 Carbone, 2013; Besson et al., 2017; Atallah-Yunes, Murphy & Noy, 2020). Only a few articles

67 report that B cells can be infected with HIV-1 (Fritsch et al., 1998; Lazzi et al., 2002; Katano et

68 al., 2007), and it seems that this is an extremely rare/artifactual event. One of the most likely

69 mechanisms of the development of HIV-1-associated B-cell lymphomas in HIV-1-infected

70 patients may be an action of viral proteins on uninfected B cells (Dolcetti et al., 2016).

71   HIV-1 produces a small nuclear transcriptional activator protein known as transactivator of

72 transcription (Tat) which regulates viral transcription (Ali et al., 2021). In the absence of Tat,

73 HIV-1 proviral promoter is incompetent for elongation: shortly after transcription initiation,

74 RNA Pol II is stalled due to the presence of inactive form of positive transcription elongation

75 factor b (P-TEFb) composed of CDK9, cyclin T1 or T2, and inhibitory 7SK small nuclear

76 ribonucleoprotein complex (containing 7SK RNA and HEXIM1) (Kao et al., 1987; Sedore et al.,

77 2007; D'Orso & Frankel, 2010). Tat is able to relieve this repressed state by binding to the TAR-

78 like sequence in 7SK snRNA and displacing HEXIM1 from cyclin T1, which disrupts the 7SK-

79 P-TEFb negative transcriptional regulatory complex and releases active P-TEFb (Barboric et al.,

80 2007; Sedore et al., 2007; Muniz et al., 2010; Pham et al., 2018) . This ability of Tat depends on

81 Tat C22 residue within the activation domain (Barboric et al., 2007). Tat then recruits the active

82 P-TEFb complex (consisting of CDK9, cyclin T1 or T2) and other transcriptional coactivators to

83 the TAR RNA element at the 5' end of stalled nascent HIV-1 transcripts to relieve stalled RNA

84 Pol II (Nekhai & Jeang, 2006; He et al., 2010). Simultaneously, HIV-1 Tat modulates cellular

85 processes by interacting with different cellular structures, particularly nuclear components

86 (Musinova et al., 2016; Ali et al., 2021).

87 Expression of HIV-1 Tat alone in mice leads to development of different neoplasms,

88 including lymphomas (Vogel et al., 1988, 1991; Corallini et al., 1993; Altavilla et al., 1999;

89 Kundu et al., 1999), suggesting that Tat protein participates in oncogenesis in HIV-1-infected

90 patients. Tat is actively released from HIV-1-infected cells (Ensoli et al., 1990; Nath, 2015) and

91 is detectable in the serum of HIV-1-infected individuals (Westendorp et al., 1995; Xiao et al.,

92 2000; Poggi et al., 2004; Germini et al., 2017). Exogenous HIV-1 Tat can enter uninfected cells,

93 and in particular, HIV-1 Tat is present within tumor cells of HIV-1-associated B-cell lymphomas

94 (Lazzi et al., 2002; Alves de Souza Rios et al., 2021). Some other HIV-1 proteins might also

95 affect cells not infectable by HIV-1, e.g., HIV-1 Nef, which can be secreted in a form of

96 extracellular vesicles and released into circulation (Raymond et al., 2011; Pushkarsky et al.,

97 2022).

98 The mechanisms of Tat-dependent lymphomagenesis in HIV-1-infected patients have been

99 described only partially. Incubation of B cells from healthy donors with recombinant HIV-1 Tat

100 *ex vivo* led to the convergence of chromosomal loci that are usually involved in t(8;14)

101 translocation, which is common in Burkitt lymphoma (Germini et al., 2017). These data indicate

102 that exogenous viral proteins can induce global rearrangement of nuclear organization and that

103 these changes can promote lymphomagenesis. Additionally, HIV-1 Tat can modify the

104 chromatin organization and gene expression of host cells, particularly T cells (Marban et al.,

105 2011; Reeder et al., 2015) and macrophages (Carvallo et al., 2017), and it seems that Tat can

106 induce a chain of similar events in B cells. Indeed, Tat can affect the expression of several genes

107 in B cells, including *AICDA*, a gene that encodes the activation-induced cytidine deaminase that

108 participates in immunoglobulin gene maturation (Sall et al., 2019; Akbay et al., 2021).

109 Additionally, HIV-1 Tat enhances *c-MYC* transcription by binding to the *c-MYC* promoter,

110 which can contribute to a more aggressive lymphoma phenotype (Lazzi et al., 2002; Alves de

111   Souza Rios et al., 2021). Thus, HIV-1 Tat present in blood may affect gene expression in B cells,

112   and these changes can promote lymphomagenesis. However, the effect of Tat on B cells has

113   never been studied on a genome-wide level. Here, we ectopically expressed HIV-1 Tat in the

114   lymphoblastoid B cell line (RPMI 8866) and analyzed host gene expression by RNA-seq. We

115   found that the expression of HIV-1 Tat led to substantial modifications of gene expression and

116   induced cellular antiviral reactions. Ectopic Tat expression also resulted in modification of

117   cellular proliferation and genome stability, thus   promoting changes that could facilitate

118   lymphomagenesis.

119

120   **MATERIALS AND METHODS**

121   **Cell culture**

122   RPMI 8866 cells (Sigma) were grown at 37°C in RPMI 1640-Gluta-Max medium (Gibco)

123   supplemented with 10% fetal bovine serum (HyClone), sodium pyruvate (PanEco), and an

124   antibiotic and antimycotic solution (Gibco).

125   HeLa cells with integrated LTR-TurboRFP (Kurnaeva et al., 2022) were grown in

126   Dulbecco's modified Eagle's medium supplemented with alanyl-glutamine (Paneco), 10% fetal

127   calf serum (HyClone) and an antibiotic and antimycotic solution (Gibco). The transactivation

128   assay based on fast-maturing TurboRFP protein was described elsewhere (Kurnaeva et al.,

129   2022). The expression of EGFP and TurboRFP was analyzed using a FACS Aria SORP

130   instrument (BD Biosciences).

131   **Plasmids and cell lines**

132   The pGST-Tat 1 86R plasmid was obtained through the NIH AIDS Reagent Program, Division

133   of AIDS, NIAID, from Dr. Andrew Rice (Herrmann & Rice, 1993, 1995; Rhim et al., 1994).

134   Plasmids for EGFP, Tat-EGFP and TatC22G-EGFP expression and lentiviral particles

135   were constructed by Evrogen (Moscow, Russia). EGFP-, Tat-EGFP- or TatC22G-EGFP-

136   expressing cells were collected using a FACSAria SORP cell sorter (BD Biosciences). The

137   excitation wavelength for EGFP was 488 nm, and the emission was detected by a 505LP and

138   515/20BP set of filters. Sorting was performed with an 85 μm nozzle and the corresponding

139   custom pressure parameters. The sorted cells were grown, then frozen in the complete medium in

140   the presence of DMSO, and stored in liquid nitrogen. To achieve better reproducibility of

141   experiments, cells were used for no more than one month after thawing (excluding the

142   experiments on long-term culture).

Cell lysate preparation, SDS-PAGE and western blotting

144   Cells were collected by centrifugation for 10 min at 800 g. Cell pellets were washed with PBS

145   and resuspended in NETN buffer (150 mM NaCl, 1 mM EDTA, 50 mM Tris pH 7.5, 0.5% NP

146   40, protease inhibitor cocktail), sonicated, incubated on ice for 30 minutes and centrifuged at 4

147   °C at 12000 g for 10 minutes. Protein quantification was performed using the Pierce™ BCA

148   Protein Assay Kit (Thermo Scientific) on a NanoDrop 2000C (Thermo Scientific). After

149   measuring the concentration, cell lysates were supplemented with Laemmli buffer and 0.1 M

150   DTT and then heated at 95°C for 10 min.

151   Protein samples (20 μg) and prestained molecular weight markers (PageRuler™ Prestained

152   Plus Protein Ladder, Thermo Scientific) were resolved on 15-well precast SDS–PAGE gels (4-

153   12%) (NuPage) in MOPS Running Buffer (NuPage). Proteins were transferred onto a 0.45 μm

154   PVDF membrane (GE Healthcare) in transfer buffer (0.025 M Tris, 0.192 M glycine, 20%

155   ethanol) at 90 V at 4°C for 2 h. Nonspecific binding was blocked in 5% nonfat dried milk in

156   Tris-buffered saline and 0.1% Tween-20 (TBST) at room temperature for 1 hour.

157   Proteins were probed at 4 °C overnight with the following primary antibodies: anti-Tat

158   (1:200, Santa Cruz, cat. #sc-65912), anti-GFP (1:1000, Roche, cat. #11814460001), anti-β-actin

159   (1:1000, control of protein load, Santa Cruz, cat. #sc-81178). The membranes were washed with

160   TBST and incubated with goat anti-mouse IgG-HRP (Santa Cruz, cat. # sc-2005) secondary

161   antibodies at a 1:2000 dilution at room temperature for 1.5 h, followed by washing in TBST.

162   Proteins were visualized using Immobilon Western Chemiluminescent HRP Substrate

163   (Millipore) and ImageQuant LAS 4000 mini (GE Healthcare) for western blotting imaging and

164   analysis.

RNA extraction and sequencing

166   Cells were collected and stored in RNALater (Qiagen). Total RNA was extracted using a

167   RNeasy Mini RNA isolation kit (Qiagen) following the Qiagen protocol, with the following

168   modifications: (1) lysis time was increased up to 40 min; (2) on-column DNase I treatment was

169   performed. RNA sample quality was assessed using a capillary electrophoresis Bioanalyzer 2100

170   (Agilent), and all samples had RIN>8. The cDNA libraries were constructed using the NEBNext

171   Ultra II Directional RNA Library Prep Kit for Illumina (NEB) following the manufacturer's

172   recommendations. RNA was fragmented for 5 minutes. Thereafter, the constructed libraries were

173   sequenced on an Illumina HiSeq 2000 with a single-end 51 bp read length. Basecalling was

174   performed using bcl2fastq v2.17.1.14.

175   **RNA-seq data processing and analysis**

176   Read   quality   control   was   performed   using   FastQC   (version   0.11.7)

177   (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Due to the good quality of the

178   reads, no filtering or adapter removal was performed. The reads were then aligned to the human

179   genome assembly GRCh38.p10 using HISAT2 (version 2.0.5) (Kim, Langmead & Salzberg,

180   2015). Read counting was performed in strand-specific mode by the htseq-count script from

181   Python library HTSeq (version 0.12.4) (Anders, Pyl & Huber, 2015) using GENCODE v26 gene

182   annotation (ALL). Genes with no counts for all samples, as well as highly expressed ribosomal

183   genes, were filtered out, resulting in the expression set of 32,120 genes across 12 samples.

184   Principal component analysis (PCA) was performed on the rlog-transformed count data,

185   and the first 2 principal components were extracted with the corresponding percentage of

186   explained variance. Differential expression analysis was performed with the R package DESeq2

187   (version 1.30.1) (Love, Huber & Anders, 2014). We declared the gene to be differentially

188   expressed if padj (p value adjusted by the Benjamini-Hochberg procedure) was smaller than 0.05

189   and the fold change was larger than 1.5 in any direction (Table S1). The statistical power for

190   protein-coding genes (with median 80 aligned reads) was calculated with R package

191   RNASeqPower (version 1.30.0) and was 0.91. Each group of samples consisted of 3 biological

192   replicates.

193   Overrepresentation analysis (ORA) and gene set enrichment analysis (GSEA)

194   (Subramanian et al., 2005) were performed with the R package clusterProfiler (version 3.18.1)

195   (Yu et al., 2012). GSEA was performed on a preranked list of genes ordered by the *stat* column

196   of DESeq2 results. The KEGG (release 99) (Kanehisa & Goto, 2000) and GO biological process

197   (Ashburner et al., 2000; The Gene Ontology Consortium, 2017) databases were used as the

198   sources of gene sets for ORA and GSEA. GO annotation was obtained from the R package

199   org.Hs.eg.db (version 3.12.0) (Marc Carlson (2020). An adjusted p value cutoff of 0.05 was used

200   to select statistically significant categories. REVIGO (Supek et al., 2011) with a cutoff parameter

201   of 0.4 was used to remove redundant GO terms.

202   To assess the possible additional activation of EBV genes due to Tat protein expression in

203   RPMI^{Tat} cells, the same pipeline for differential expression analyses was used. The EBV gene

204   annotation file in GTF format was obtained from GenBank file NC_007605.1 using a custom

205   Python script. Raw reads were aligned to combined human and viral genomes. The EBV GTF

206   annotation file was used to obtain counts for viral genes, and previously obtained human gene

207   counts were used to estimate size factors for all samples. Subsequent differential expression

208   analysis was performed for viral genes.

209   **RNA extraction and qRT-PCR assays**

210   Total RNA from RPMI 8866 cells was isolated using the RNeasy Mini Kit (Qiagen). The RNA

211   concentration was measured with a NanoPhotometer (Implen). Reverse transcription was

212   performed with an iScript Advanced cDNA Synthesis Kit (BioRad) according to the

213   manufacturer's instructions, and qPCR was performed in technical triplicates using a SYBR

214   Green kit (Syntol) in a CFX96 Real-Time PCR Detection System (BioRad). Melting curve

215   analyses were performed to verify the amplification specificity. Experiments were performed in

216   biological triplicates, and error bars represent the SEM as indicated in all figure legends. The

217   HPRT, YWHAZ and UBC2 genes were used as references. The primers used for qRT-PCR

218   analysis are listed in Table S2.

219   **Electron microscopy**

220   The cells were fixed in 4.0% glutaraldehyde in 0.1 M cacodylate buffer for 8 h, postfixed with

221   1% osmium tetroxide for 1 h, dehydrated in ethanol and acetone (70% ethanol containing 2%

222   uranyl acetate), and embedded in Spi-pon 812 epoxy resin (SPI Inc.). Ultrathin sections were cut

223   using an Ultracut E Ultratome (Reichert Jung), stained with lead citrate, and photographed using

224   a JEM-1400 electron microscope (Jeol).

225   **Analysis of the cell cycle**

226   Cells were incubated in the presence of 1 μg/ml EdU for 15 min at 37°C, washed in PBS, fixed

227   with 3.7% paraformaldehyde for 10 min and permeabilized in 0.5% Triton X-100. EdU was

228   labeled using a Click-iT EdU Cell Proliferation Kit for Imaging, Alexa Fluor 555 dye (Thermo),

229   according to the manufacturer's instructions. DNA was stained with 1 μg/ml Hoechst 33342

230   (Thermo). Cells were analyzed using a FACSAria SORP cell sorter (BD Biosciences). The

231   detection parameters were as follows: Ex. 405 nm, Em. 450/50 BP for Hoechst 33342 and Ex.

232   561 nm, Em. 585/15 BP for EdU-Alexa Fluor 555.

233   For Ki-67 staining, cells were fixed with 1% paraformaldehyde for 10 min, washed in PBS,

234   permeabilized in 0.01% Triton X-100 for 10 min, washed in PBS and stained with anti-Ki-67

235   PE-conjugated antibodies (BD Pharmingen) according to the manufacturer's instructions. Cells

236   were analyzed using a FACSAria SORP cell sorter (BD Biosciences). PE fluorescence was

237   detected at Ex. 561 nm, Em. 585/15 BP.

238   **Analysis of apoptosis**

239   Cell death was analyzed by flow cytometry using a FACSAria SORP instrument (BD

240   Biosciences). Cells were simultaneously stained with 1 µg/ml Hoechst 33342 (Thermo), 100 nM

241   TMRE (tetramethylrhodamine, ethyl ester, perchlorate, Thermo) and annexin V-Alexa Fluor 647

242   (Thermo) according to the manufacturer's instructions. Thus, the DNA content, mitochondrial

243   membrane potential and phosphatidylserine externalization were analyzed together for each

244   sample. The detection parameters were as follows: Ex. 405 nm, Em. 450/50 BP for Hoechst

245   33342; Ex. 561 nm, Em. 585/15 BP for TMRE and Ex. 647 nm, Em. 640/14 BP for annexin V-

246   Alexa Fluor 647. Additionally, the cells were stained with anti-caspase 3 PE-conjugated

247   antibodies (PE active caspase-3 apoptosis kit, BD Biosciences). Cell fixation, permeabilization

248   and staining were performed according to the manufacturer's instructions. Active caspase 3-PE

249   was detected at Ex. 561 nm, Em. 585/15 BP using a FACSAria SORP cell sorter (BD

250   Biosciences).

251   **Chromosome preparations, FISH and cytogenetic analysis**

252   For metaphase chromosome preparations, colcemid (Invitrogen) was added to cultivation media

253   at a final concentration of 0.1 µg/ml. Cells were collected 3 hours after the addition of colcemid.

254   Hypotonic treatment (0.075 M KCl) was performed for 15 min at 37°C. Cells were fixed with an

255   ice-cold mix of methanol and glacial acetic acid (3:1). Metaphase slides were made according to

256   standard procedures and used for FISH one day after preparation. FISH was performed according

257   to the manufacturer's recommendation with a mix of DNA probes specific to whole human

258   chromosomes 1, 2 and 4 (Metasystems, Germany). The DNA probes to human chromosomes 1,

259   2 and 4 were labeled by fluorochromes of red, green and both colors, respectively.

260   Metaphase identification and image acquisition were performed with the slide scanning

261   platform Metafer (v.3.11.8, Metasystems) installed on an upright light microscope (Axioscope

262   A1, Carl Zeiss, Germany). For image processing, the Isis FISH imaging system (v5.5, release

263 5.5.10, Metasystems) was used. The number of metaphases scored per sample in each replicate
264 varied from 1091 to 1592.

265   We used Fisher's exact test to estimate the statistical significance of differences in the
266 level of chromosomal aberrations. Differences were considered statistically significant at a
267 significance level of p <0.01.

268

269 **RESULTS**

270 **Generation of cell lines for the analysis of HIV-1 Tat action in B cells**

271 To analyze the effect of HIV-1 Tat on B cells, we developed RPMI 8866-based cell lines stably
272 expressing Tat protein fused with EGFP (hereafter referred to as RPMI^{Tat}) (Fig. 1A). As controls,
273 we constructed cell lines expressing either EGFP (RPMI^{EGFP}) or TatC22G-EGFP, a mutant Tat
274 protein deprived of transactivation capacity (RPMI^{Cys}).

275   Stable lines were obtained by transduction of cells with pseudoviral particles, and cells
276 expressing the proteins of interest were selected using a fluorescence-activated sorter (FACS)
277 (Fig. 1B). Flow cytometry demonstrated high purity and homogeneity of the obtained cell lines.
278 Of note, there was an admixture of nonfluorescent cells in the RPMI^{Tat} and RPMI^{Cys} cell lines.
279 Tat-EGFP and TatC22G-EGFP expression was confirmed by western blotting (Fig. 1C). Tat-
280 EGFP and TatC22G-EGFP were partially proteolysed, and as a result, additional bands at ~27
281 kDa were visible (nonprocessed images are presented in Figure S1). As the EGFP tag could have
282 interfered with Tat activity, we analyzed the transactivation capacity of Tat-EGFP and TatC22G-
283 EGFP using an *in vitro* assay based on the fast-maturing fluorescent protein TurboRFP.
284 TurboRFP expression in HeLa cells was controlled by a fragment of the HIV-1 3' LTR
285 (Kurnaeva et al., 2022). We transduced EGFP, Tat-EGFP, or TatC22G-EGFP into these HeLa
286 cells and found that the expression of Tat-EGFP substantially increased TurboRFP fluorescence
287 compared to EGFP, which did not cause an increase in TurboRFP expression (Fig. 1D).
288 TatC22G had a ~17 fold decreased transactivation activity as compared to Tat (the median
289 fluorescence intensity of TurboRFP in Tat-expressing cells was 34601, the median fluorescence
290 intensity in TatC22G-expressing cells was 2009, and the median fluorescence intensity in EGFP-
291 expressing cells was 309 in the representative experiment shown in Figure 1D). Hence, the
292 transactivation activity of the Tat protein was not perturbed by its fusion with EGFP.

293 **HIV Tat protein affects the gene expression profile of RPMI 8866 cells**

294   To determine the genes regulated by HIV-1 Tat, total RNA from RPMI 8866, RPMI^{EGFP},

295   RPMI^{Tat} and RPMI^{Cys} cell lines was collected, and RNA-seq was performed. Three biological

296   replicates were sequenced for each cell line. From 16 to 24 million 51 nt sequencing reads were

297   generated for the RPMI^{Tat}, RPMI^{Cys}, RPMI^{EGFP} and RPMI samples. On average, 67.17% of reads

298   in every sample were uniquely aligned to the reference genome GRCh38, and 78.93% of them

299   were nonambiguously mapped to the GENCODE gene annotation (Fig. S2). To assess variability

300   between replicates and between cell lines, rlog-transformed filtered count data were visualized

301   by principal component analysis (Fig. S3A). Biological replicates proved to be highly alike by

302   clustering tightly according to sample type. RPMI^{Tat} samples tended to cluster with RPMI^{Cys}

303   samples, implying similar effects of Tat-EGFP and TatC22G-EGFP on gene expression. This

304   finding was also confirmed by Spearman correlation of normalized gene expression profiles (Fig.

305   S3B).

306   To identify changes in gene expression induced by Tat protein in RPMI 8866 cells, we

307   performed differential expression analysis for each cell line. We qualified differentially

308   expressed genes (DEGs) with padj < 0.05 and fold change >1.5 in any direction.

309   Comparison of the control RPMI^{EGFP} against the RPMI cell line demonstrated that the

310   impact of EGFP on gene expression could be neglected, as the analysis revealed only 17

311   upregulated and 36 downregulated DEGs, implying that the gene expression profiles of

312   RPMI^{EGFP} and RPMI samples were highly similar (Fig. 2A). Therefore, we further used RPMI

313   cells as the control and their gene expression level as the baseline to identify genes regulated by

314   Tat or TatC22G proteins.

315   Out of 32,120 genes with detectable expression, 1038 genes were differentially expressed

316   between RPMI^{Tat} and RPMI cells. A comparable number of DEGs (902 genes) were identified in

317   the comparison RPMI^{Cys} versus the control (Fig. 2A). Among the detected DEGs, slightly more

318   genes were downregulated (60 or 55%) than upregulated in the presence of Tat or TatC22G

319   protein, respectively. The majority of genes whose expression was affected were protein-coding

320   genes. We validated the expression of several DEGs using quantitative real-time polymerase

321   chain reaction (qRT-PCR) assays, thus confirming the reliability of RNA-seq (Fig. 2B).

322   RPMI 8866 is an EBV-positive B-lymphoblastoid cell line derived from a patient with

323   chronic myelogenous leukemia (McCune, Fu & Kunkel, 1981). RPMI 8866 cells express

324   EBNA1, -2, -3A, -3B, -3C, and LMP-1, -2A, and -2B proteins and several noncoding RNAs. To

325   explore the possibility that changes in gene expression were a result of Tat-induced changes in

326   the expression of EBV genes, we performed differential gene expression analysis of the EBV

327   transcriptome and found no evidence of Tat or TatC22G protein impact on the expression of

328   viral genes in RPMI^{Tat} or RPMI^{Cys} cells (Fig. S4); thus we concluded that we observed a direct

329   effect of Tat or TatC22G on the host cell.

330   Next, we performed gene set enrichment analysis (GSEA) and overrepresentation analysis

331   (ORA) to search for activated or suppressed functional gene groups and molecular pathways as

332   defined by the GO Biological Process (GO BP) and KEGG databases.

333   The analysis of modified biological processes (GO BP) was performed by GSEA with

334   subsequent removal of redundant GO terms with REVIGO (Fig. 2C). Comparison of RPMI^{Tat}

335   cells with the control RPMI cells showed that genes whose expression was upregulated in the

336   presence of Tat protein were enriched for antiviral responses, including *Regulation of defense*

337   *response to virus* (p.adjust = 4.3×10^{-2}), *Type I interferon production* (p.adjust = 1.9×10^{-2}), *Type I*

338   *interferon signaling pathway* (p.adjust = 1.1×10^{-3}), *Viral genome replication* (p.adjust = 1.4×10^{-2}

339   ), *Regulation of viral process* (p.adjust = 1.6×10^{-2}), and *Negative regulation of viral processes*

340   (p.adjust = 9.2×10^{-3}) (Fig. 2C, left panel). In addition to the activation of pathways associated

341   with antiviral responses, the activation of *Regulation of DNA-template transcription, elongation*

342   (p.adjust = 5.9×10^{-3}), *mRNA processing* (p.adjust = 2.2×10^{-3}), *Regulation of mRNA processing*

343   (p.adjust = 4.3×10^{-2}), and *G1/S phase transition of mitotic cell cycle* (p.adjust = 2.5×10^{-2}) was

344   also observed in RPMI^{Tat} cells. These changes may be either part of the global antiviral response

345   or the result of the action of a viral protein (proviral reaction). Among the downregulated

346   pathways, pathways associated with cell adhesion (*Adherens junction organization* (p.adjust =

347   3.1×10^{-2}), *Cell-cell adhesion* (p.adjust = 2.6×10^{-6}), *Positive regulation of cell adhesion* (p.adjust

348   = 2.5×10^{-2})), and multiple metabolic and biosynthetic pathways were detected (Fig. 2C, right

349   panel). Notably, genes involved in the regulation of cellular proliferation were downregulated

350   (*Regulation of leukocyte proliferation* (p.adjust = 2.5×10^{-2}), *Lymphocyte proliferation* (p.adjust =

351   8.4×10^{-3})).

352   We additionally analyzed changes in signaling pathways, for which overrepresentation

353   analysis (ORA) using KEGG databases was performed (Fig. 3). Analysis of up- and

354   downregulated DEGs revealed activation of innate immunity pathways, such as *Toll-like*

355   *receptor signaling pathways* (hsa04620; p.adjust = 3.1×10^{-2}) and *NOD-like receptor signaling*

356   *pathways* (hsa04621; p.adjust = 3.1×10^{-2}), as well as the *Influenza A* (hsa05164; p.adjust =

357   3.1×10^{-2}) pathway, which was highly overlapping with the former two. *Herpes simplex virus 1*

358   *infection* pathway (hsa05168; p.adjust = 3.1×10^{-2}) seems to be related to the antiviral response

359   (Fig. 3A). Several zinc finger proteins (ZNFs) were annotated as related to this pathway.

360   Notably, several genes of the JAK-STAT signaling pathway (*JAK1* and *STAT1*) were general for

361   all detected pathways (Fig. 3B). Secreted interferon-I induces IFN-stimulated genes through the

362   JAK-STAT pathway (Ivashkiv & Donlin, 2014). Some ZNFs regulate the transcription of

363   interferon-stimulated genes (Wang & Zheng, 2021), but among ZNF genes detected by RNA-

364   seq, there was only one well-investigated protein: ZNF268 (Table S3). ZNF268 was described as

365   chronic lymphocytic leukemia (CLL)-associated antigen (Krackhardt et al., 2002), its aberrant

366   alternative splicing was detected in human hematological malignancies (Zhao et al., 2008).

367   ZNF268 also contributes to cervical carcinogenesis through the NF-κB signaling pathway (Wang

368   et al., 2012).

369   Downregulated DEGs were enriched for *Cytokine-cytokine receptor interaction* pathway

370   (hsa04060; p.adjust = 2.0×10^{-6}) and overlapped with *Viral protein interaction with cytokine and*

371   *cytokine receptor* pathway (hsa04061; p.adjust = 2.9×10^{-5}), reflecting possible inhibition of the

372   proinflammatory response (Fig. 3C). The downregulated genes of these pathways included

373   several cytokines (*IL16*, *IL19*, *TNFSF9*, *CCL20*, *CCL22*, *CFS1*, etc.) and receptors of cytokines

374   (*IL1R1*, *IL6R*, *IL7R*, *CCR4*, *CCR8*, *TNFRSF19*, *TNFRSF11A*, etc.) (Fig. 3D). At the same time,

375   some proinflammatory cytokines were upregulated (for example, *IL6*, FC 2.12). Additionally,

376   pathways associated with *Cell adhesion molecules* (hsa04514; p.adjust = 8.0×10^{-7}), and

377   *Leukocyte transendothelial migration* (hsa04670; p.adjust = 2.2×10^{-2}) were also suppressed in

378   RPMI^{Tat} cells, indicating possible inhibition of cell adhesion (Fig. 3C). Among the genes of *Cell*

379   *adhesion molecules* pathway, several genes (*HLA-DQA1*, *HLA-DQB1*, *HLA-DRB1*, and *HLA-*

380   *DRB5*) coding an MHC class II cell surface receptor were suppressed, which might affect the T

381   cell receptor signaling pathway (Fig. 3D).

382   Additionally, we analyzed the expression of different transcription factors because such

383   changes can lead to pronounced effects. We found that the expression of 12 transcription factors

384   was upregulated (*ASCL1*, *ATF5*, *FOXO3*, *NR2F2*, *PBX1*, *PRDM5*, *REL*, *SIX1*, *STAT1*, *TFCP2*,

385   *ZNF268*, *ZNF740*), and 13 were downregulated (*DMRT2*, *FLI1*, *HOXB7*, *JUN*, *KDM3A*, *MYCN*,

386   *NR1H4*, *SOX5*, *SOX6*, *SSX1*, *ZC3H6*, *ZNF358*, *ZNF613*); thus, the action of Tat (direct or

387 indirect) may be at least partially connected with the regulation of transcription factor
388 expression.

389   In addition to differentially expressed protein coding genes, the expression levels of 154
390 long non-protein coding RNAs (lncRNAs) were altered by HIV-1 Tat expression (54 and 100
391 lncRNAs were upregulated and downregulated, respectively). Differentially expressed lncRNAs
392 accounted for 14.8% of all DEGs between RPMI^{Tat} and control cells. These lncRNAs are poorly
393 characterized, but two genes of lncRNAs among the identified DEGs are well known. The
394 MALAT1 lncRNA involved in transcriptional regulation and alternative splicing (West et al.,
395 2014) was upregulated 2.44-fold in RPMI^{Tat} over the control (the elevated expression was
396 additionally confirmed by qRT-PCR, Fig. S5A). MALAT1 is upregulated in HIV-1-infected
397 CD4$^+$ T cells (Qu et al., 2019) and in the peripheral blood of HIV-1-infected patients (Jin et al.,
398 2016). NEAT1 lncRNA forms the core structural component of paraspeckle bodies and is known
399 for its contribution to HIV replication (Zhang et al., 2013). We observed a 1.66-fold increase in
400 *NEAT1* expression in the RPMI^{Tat} B cell line (see also Fig. S5B), indicating that NEAT1 may
401 play an important role in the development of cell responses induced by the Tat protein.

402   We also analyzed other DE lncRNAs. The complete list of DE lncRNAs is presented in
403 Table S4, and, since the level of some of these lncRNAs was extremely low, we additionally
404 analyzed ten most expressed DE lncRNAs (Table 1). Some of DE lncRNAs, which were
405 differentially expressed in RPMI^{Tat} cells, are also differentially expressed in different tumors and
406 might be involved in development of these tumors. The most known examples are *MALAT1* and
407 *NEAT1* expressed in different neoplasms, including B-cell lymphomas and leukemias (Table S5).

408 **Comparison of Tat- and TatC22G-induced effects**

409 HIV-1 Tat can modify host gene expression by several nonrelated mechanisms. Tat can
410 transactivate host genes by binding to TAR-like sequences in their nascent mRNAs, e.g., IL6
411 mRNA (Ambrosino et al., 1997) or TNF-β mRNA (Buonaguro et al., 1994). Tat can also
412 associate with chromatin and control RNA polymerase II recruitment and pause release (Reeder
413 et al., 2015). Additionally, HIV-1 Tat interacts with hundreds of nuclear proteins (Gautier et al.,
414 2009), and this interaction may also modify the expression of host genes.

415   Transactivation of transcription by the mutant TatC22G protein was significantly weaker
416 than that by Tat (Fig. 1D), but it could still bind to chromatin and interact with other proteins.
417 Both Tat and TatC22G proteins modulate the expression of similar and overlapping sets of genes

418 (Fig. 4A). Comparison of the transcriptome profiles of the RPMI^{Tat} and RPMI^{Cys} cell lines

419 revealed 201 protein-coding DEGs (padj < 0.05; fold change >1.5 in any direction).

420 Overrepresentation analysis against the KEGG database also demonstrated little gene set

421 enrichment, reflecting relatively small differences between RPMI^{Cys} and RPMI^{Tat} cells (Fig. 4B).

422 Notably, the *Cytokine-cytokine receptor interaction* pathway (hsa04060) downregulated by HIV-

423 1 Tat was further suppressed by mutant TatC22G protein in RPMI^{Cys} cells compared to RPMI^{Tat}

424 cells.

425

426 **The pathological effects of Tat expression**

427 Our RNA-seq analysis identified several pathways affected by Tat. We next experimentally

428 analyzed key characteristics of RPMI 8866, RPMI^{EGFP}, RPMI^{Tat}, and RPMI^{Cys} cell lines to

429 identify how Tat affected B cells.

430 Electron microscopy demonstrated that the expression of EGFP, Tat-EGFP or TatC22G-

431 EGFP did not lead to any substantial changes in cellular organization (Figure 5A). To ascertain

432 whether changes in the expression of several genes involved in the G_1/S phase transition and cell

433 proliferation (Fig. 2C) affected cell cycle progression in Tat-expressing cells, we analyzed the

434 proportion of G_1/G_0, S, and G_2/M cells in the total population of RPMI^{Tat} cells using flow

435 cytometry. We could not detect any significant changes in the proportions of cells at different

436 stages of the cell cycle using this approach (Figs. 5B, C). We also detected S-phase cells using

437 incorporation of EdU and did not observe any difference between different cell lines (Fig. 5D).

438 Finally, we analyzed fractions of cycling (non-G_0) cells using antibodies against the proliferation

439 marker Ki-67, and again, no difference was found (Fig. 5E).

440 Cell death can also influence the dynamics of cell populations; therefore we analyzed

441 apoptosis in RPMI^{Tat} cells. We did not observe any statistically significant changes in the content

442 of sub-G_1 cells (dead cells), cells with reduced mitochondrial potential (TMRE staining, which

443 allowed for marker cells during very early stages of apoptosis), proportion of annexin V-positive,

444 or caspase 3-positive cells (markers of apoptotic cells) (Fig. 5F).

445 Thus, stable expression of Tat did not produce strong effects that could be detected using

446 standard analysis of cellular proliferation and apoptosis. At the same time, subtle, but persistent

447 changes (if they exist) could potentially influence the dynamics of cell growth. Therefore, we

448 cultured cells for three months and found that the percentage of cells expressing Tat-EGFP was

449 gradually reduced (Figs. 6A, B). Such replacement of Tat-expressing cells by non-expressing

450 cells was not detected in cells expressing either TatC22G-EGFP or EGFP, clearly indicating that

451 this effect was a consequence of Tat protein *per se*.

452   To detect possible mutagenic effects of HIV-1 Tat expression, we performed cytogenetic

453 analysis. We scored chromosome aberrations involving chromosomes 1, 2 and 4 revealed by

454 fluorescence *in situ* hybridization with whole chromosome paints (Fig. 6C). The images of

455 metaphases were automatically acquired, and then metaphases were analyzed for the presence of

456 structural chromosomal aberrations, including translocations, dicentrics, acentrics, and deletions.

457 Numerical chromosomal abnormalities, such as polyploidy and aneuploidy, were also taken into

458 consideration. We observed a reproducible increase in the frequency of translocations as well as

459 in the total yield of structural aberrations in cells expressing either Tat-EGFP or TatC22G-EGFP

460 (Fig. 6D; Table S6).

461

462 **Discussion**

463 To identify the mechanisms potentially leading to the development of B-cell lymphomas in HIV-

464 infected patients, we investigated the effect of Tat protein on gene expression in cultured B cells.

465 Upon long-term exposure of B-cells to Tat, the effects could accumulate and potentially provoke

466 an oncogenic transformation. Therefore, we did not use either an *ex vivo* experimental model, in

467 which lymphocytes isolated from the blood of healthy donors were incubated in the presence of

468 Tat protein (for example, (Germini et al., 2017)), or a model with the expression of HIV-1

469 controlled by an inducible promoter (for example, (Reeder et al., 2015)). These experimental

470 systems are more likely to simulate an acute infection situation, but they do not allow for the

471 study of weak long-term effects of protein and compensatory reactions of the cell. In the current

472 study, we obtained the B lymphoid cell line RPMI 8866 stably expressing HIV-1 Tat fused with

473 EGFP and used these cells to simulate a prolonged systemic effect of the presence of Tat in

474 blood of chronically infected patients.

475   To detect possible changes provoked by chronic exposure of B-cells to Tat, we performed

476 a genome-wide analysis of cellular gene expression (RNA-seq). Tat protein induces differential

477 expression of approximately 1000 genes (padj < 0.05 and fold change of ≥ 1.5 in any direction).

478 To predict the effects of these changes, the analysis of metabolic and signaling pathways (GO BP

479 and KEGG analysis) was performed. GO BP analysis demonstrated the activation of pathways

480 involved in cellular antiviral reactions, and suppression of different metabolic pathways and
481 proliferation. KEGG analysis also demonstrated activation of innate immunity pathways
482 involved in antiviral reactions: *Toll-like receptor signaling* and *NOD-like receptor signaling*
483 pathways. We observed an increase in *JAK1* and *STAT1* expression, indicating that Tat may
484 affect these pathways *via* the JAK-STAT signaling. Additionally, we found the reduction of
485 Cytokine-cytokine receptor interaction pathway. Downregulation of this pathway (and
486 downregulation of metabolic pathways and proliferation) may result from direct HIV-1 Tat
487 action (proviral reactions). It should be noted however that not all cytokines and their receptors
488 were downregulated. HIV-1 Tat induces the expression of several proinflammatory cytokines,
489 mainly IL-6 in several cell types (Ambrosino et al., 1997; Nookala & Kumar, 2014; Ben Haij et
490 al., 2015). Our results also confirm the upregulation of pro-inflammatory IL6 gene in Tat-
491 expressing B cells. Thus, the gene expression pattern we observed in Tat-expressing B cells
492 probably resulted from a combination of Tat action per se and cellular antiviral reactions.

493   We also analyzed the lncRNA expression. Differentially expressed lncRNAs accounted for
494 14.8% of all DEGs between RPMI^{Tat} and the control cells. While functions of the majority of
495 identified lncRNA are poorly studied, some of them (e.g., NEAT1 and MALAT1) play a role in
496 development of B-cell neoplasms (Table S5). In Tat-expressing RPMI 8866 cells, both of these
497 lncRNAs were upregulated. It is possible that these or some other DE lncRNAs can be involved
498 in the development of B-cell lymphomas in HIV-1-infected patients, but this requires further
499 study.

500   Another important observation on the mode of action of Tat in host cells came from the
501 comparison of the action of Tat and its TatC22G mutant with the reduced transactivation
502 activity. In our experiments, 313 protein-coding genes were regulated by Tat only (115
503 upregulated and 198 downregulated, respectively), 228 genes were regulated by TatC22G only
504 (117 upregulated and 111 downregulated genes), while the most genes (435) were regulated both
505 by Tat and TatC22G (185 upregulated and 250 downregulated genes) (Figure 4). HIV-1 Tat can
506 modify cellular processes *via* different mechanisms. It can transactivate host genes by binding to
507 TAR-like sequences (for example, this mechanism seems to regulate the transcription of *IL6* in T
508 cells (Ambrosino et al., 1997) and *TNF-b* in B cells (Buonaguro et al., 1994)). Tat can also
509 interact with chromatin and RNA-polymerase II (Reeder et al., 2015) as well as with hundreds of
510 other cellular and nuclear proteins (Gautier et al., 2009). Therefore, to distinguish between the

511   two potential modes of action of Tat, we compared the effects of HIV-1 Tat and that with the

512   C22G mutation possessing a significantly decreased transactivation activity. We hypothesized

513   that genes differentially expressed between RPMI^{Tat} and RPMI^{Cys} cells would be associated with

514   the transactivator activity of this protein while genes similarly regulated would rather be affected

515   by protein-protein interactions involving Tat. The RNA-seq results demonstrated that most of the

516   effects of EGFP-Tat and EGFP-TatC22G were similar, and hence, the main mechanism of

517   prolonged Tat action on B-cells seems to be due to Tat interaction with host proteins.

518   We next analyzed some physiological parameters of RPMI^{Tat} cells. We could not find any

519   manifested changes in cellular morphology, proliferation, or apoptosis. At the same time, we

520   observed an effect of prolonged cultivation, i.e., a decrease in EGFP-Tat-expressing cells,

521   indicating that Tat inhibited cell growth. Thus, although there were no strong effects of EGFP-

522   Tat expression, a weak effect, which manifested itself only in situations of long observation, was

523   induced. Additionally, we found that chromosome aberrations occurred more frequently in cell

524   lines expressing EGFP-Tat and EGFP-TatC22G. These observations are consistent with

525   published data obtained using an *ex vivo* model (El-Amine et al., 2018). Modification of cell

526   cycle progression and chromosome aberrations may both promote lymphomagenesis.

527   Thus, the presence of HIV-1 Tat can in the long run modify the cellular physiology and

528   genome stability of cultured B cells and, as a result, may promote oncogenic transformation. The

529   precise mechanisms of these effects will be the subject of our future work.

530

531   **Acknowledgement**

532   We are grateful to A.V. Lazarev for technical support. The following reagent was obtained

533   through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: HIV-1 HXB2 GST-

534   Tat Expression Vector (GST-Tat 1 (86R)) from Dr. Andrew Rice. This work was carried out

535   within the framework of the International Research Network (IRN) ONCO3D and the IDB RAS

536   Government basic research program (0088-2021-0007). The flow cytometry facility became

537   available in the framework of the Moscow State University Development Program.

538

539   **ADDITIONAL INFORMATION AND DECLARATIONS**

540   **Funding**

541 The work was supported by the Russian Science Foundation (grant 21-74-20134 to E.V.S.) and
542 the Russian Foundation for Basic Research (PhD student grant 20-34-90156 to M.A.T.). The
543 funders had no role in study design, data collection and analysis, decision to publish, or
544 preparation of the manuscript.
545

546 **Grant Disclosures**

547 The following grant information was disclosed by the authors:
548 Russian Science Foundation: 21-74-20134.
549 Russian Foundation for Basic Research: 20-34-90156.
550

551 **Competing Interests**

552 The authors declare that they have no competing interests.

553

554 **Data Availability**

555 The following information was supplied regarding data availability:

556 Raw RNA-seq data as well as processed data (gene counts) are deposited in GEO with an
557 accession number GSE182538. The code is publicly available at
558 https://github.com/kirushka/Tat_in_B_cells. The raw data are also available in the Supplemental
559 Files.

560

561 **REFERENCES**

562 **Akbay B, Germini D, Bissenbaev AK, Musinova YR, Sheval EV, Vassetzky Y,**
563 **Dokudovskaya S. 2021.** HIV-1 Tat Activates Akt/mTORC1 Pathway and AICDA
564 Expression by Downregulating Its Transcriptional Inhibitors in B Cells. *International*
565 *journal of molecular sciences* 22. DOI: 10.3390/ijms22041588.
566 **Ali A, Mishra R, Kaur H, Chandra Banerjea A. 2021.** HIV-1 Tat: An update on
567 transcriptional and non-transcriptional functions. *Biochimie* 190:24–35. DOI:
568 10.1016/j.biochi.2021.07.001.
569 **Altavilla G, Trabanelli C, Merlin M, Caputo A, Lanfredi M, Barbanti-Brodano G,**
570 **Corallini A. 1999.** Morphological, histochemical, immunohistochemical, and ultrastructural
571 characterization of tumors and dysplastic and non-neoplastic lesions arising in BK virus/tat

572   transgenic mice. *The American journal of pathology* 154:1231–1244. DOI: 10.1016/S0002-

573   9440(10)65375-8.

574   **Alves de Souza Rios L, Mapekula L, Mdletshe N, Chetty D, Mowla S. 2021.** HIV-1

575   transactivator of transcription (Tat) co-operates with AP-1 factors to enhance c-MYC

576   transcription. *Frontiers in cell and developmental biology* 9. DOI:

577   10.3389/fcell.2021.693706.

578   **Ambrosino C, Ruocco MR, Chen X, Mallardo M, Baudi F, Trematerra S, Quinto I, Venuta**

579   **S, Scala G. 1997.** HIV-1 Tat induces the expression of the interleukin-6 (IL6) gene by

580   binding to the IL6 leader RNA and by interacting with CAAT enhancer-binding protein

581   beta (NF-IL6) transcription factors. *The Journal of biological chemistry* 272:14883–14892.

582   DOI: 10.1074/jbc.272.23.14883.

583   **Anders S, Pyl PT, Huber W. 2015.** HTSeq--a Python framework to work with high-throughput

584   sequencing data. *Bioinformatics* 31:166–169. DOI: 10.1093/bioinformatics/btu638.

585   **Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K,**

586   **Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S,**

587   **Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000.** Gene ontology:

588   tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics* 25:25–

589   29. DOI: 10.1038/75556.

590   **Atallah-Yunes SA, Murphy DJ, Noy A. 2020.** HIV-associated Burkitt lymphoma. *The Lancet.*

591   *Haematology* 7:e594–e600. DOI: 10.1016/S2352-3026(20)30126-5.

592   **Autran B, Carcelain G, Li TS, Blanc C, Mathez D, Tubiana R, Katlama C, Debré P,**

593   **Leibowitch J. 1997.** Positive effects of combined antiretroviral therapy on CD4+ T cell

594   homeostasis and function in advanced HIV disease. *Science* 277:112–116. DOI:

595   10.1126/science.277.5322.112.

596   **Barboric M, Yik JHN, Czudnochowski N, Yang Z, Chen R, Contreras X, Geyer M, Matija**

597   **Peterlin B, Zhou Q. 2007.** Tat competes with HEXIM1 to increase the active pool of P-

598   TEFb for HIV-1 transcription. *Nucleic acids research* **35:**2003–2012. DOI:

599   10.1093/nar/gkm063.

600   **Ben Haij N, Planès R, Leghmari K, Serrero M, Delobel P, Izopet J, BenMohamed L,**

601   **Bahraoui E. 2015.** HIV-1 Tat Protein Induces Production of Proinflammatory Cytokines by

602   Human Dendritic Cells and Monocytes/Macrophages through Engagement of TLR4-MD2-

603 CD14 Complex and Activation of NF-κB Pathway. *PloS one* 10:e0129425. DOI:

604 10.1371/journal.pone.0129425.

605 **Besson C, Lancar R, Prevot S, Algarte-Genin M, Delobel P, Bonnet F, Meyohas M-C,**

606 **Partisani M, Oberic L, Gabarre J, Goujard C, Boue F, Coppo P, Costello R, Hendel-**

607 **Chavez H, Mekerri N, Dos Santos G, Recher C, Delarue R, Casasnovas R-O, Taoufik**

608 **Y, Mounier N, Costagliola D, ANRS-CO16 LYMPHOVIR Cohort. 2017.** Outcomes for

609 HIV-associated diffuse large B-cell lymphoma in the modern combined antiretroviral

610 therapy era. *AIDS* 31:2493–2501. DOI: 10.1097/QAD.0000000000001652.

611 **Buonaguro L, Buonaguro FM, Giraldo G, Ensoli B. 1994.** The human immunodeficiency

612 virus type 1 Tat protein transactivates tumor necrosis factor beta gene expression through a

613 TAR-like structure. *Journal of virology* 68:2677–2682. DOI: 10.1128/JVI.68.4.2677-

614 2682.1994.

615 **Carvallo L, Lopez L, Fajardo JE, Jaureguiberry-Bravo M, Fiser A, Berman JW. 2017.**

616 HIV-Tat regulates macrophage gene expression in the context of neuroAIDS. *PloS one*

617 12:e0179882. DOI: 10.1371/journal.pone.0179882.

618 **Chen Q, Shen H, Zhu X, Liu Y, Yang H, Chen H, Xiong S, Chi H, Xu W. 2020.** A nuclear

619 lncRNA Linc00839 as a Myc target to promote breast cancer chemoresistance via

620 PI3K/AKT signaling pathway. *Cancer science* **111:**3279–3291. DOI: 10.1111/cas.14555.

621 **Chong C, Müller M, Pak H, Harnett D, Huber F, Grun D, Leleu M, Auger A, Arnaud M,**

622 **Stevenson BJ, Michaux J, Bilic I, Hirsekorn A, Calviello L, Simó-Riudalbas L, Planet**

623 **E, Lubiński J, Bryśkiewicz M, Wiznerowicz M, Xenarios I, Zhang L, Trono D, Harari**

624 **A, Ohler U, Coukos G, Bassani-Sternberg M. 2020.** Integrated proteogenomic deep

625 sequencing and analytics accurately identify non-canonical peptides in tumor

626 immunopeptidomes. *Nature communications* **11:**1293. DOI: 10.1038/s41467-020-14968-9.

627 **Corallini A, Altavilla G, Pozzi L, Bignozzi F, Negrini M, Rimessi P, Gualandi F, Barbanti-**

628 **Brodano G. 1993.** Systemic expression of HIV-1 tat gene in transgenic mice induces

629 endothelial proliferation and tumors of different histotypes. *Cancer research* 53:5569–5575.

630 **Cozzi E, Neddermeyer A, Miliara S, Lennartsson A, Lehmann S. 2022.** Identification of long

631 non-coding RNAs involved in leukemogenesis and venetoclax response in acute myeloid

632 leukemia through functional CRISPR-dCas9 interference screens. Available from:

633 http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-469815

634 **Dolcetti R, Gloghini A, Caruso A, Carbone A. 2016.** A lymphomagenic role for HIV beyond
635   immune suppression? *Blood* 127:1403–1409. DOI: 10.1182/blood-2015-11-681411.

636 **D'Orso I, Frankel AD. 2010.** RNA-mediated displacement of an inhibitory snRNP complex
637   activates transcription elongation. *Nature structural & molecular biology* **17:**815–821. DOI:
638   10.1038/nsmb.1827.

639 **El-Amine R, Germini D, Zakharova VV, Tsfasman T, Sheval EV, Louzada RAN, Dupuy C,**
640   **Bilhou-Nabera C, Hamade A, Najjar F, Oksenhendler E, Lipinski M, Chernyak BV,**
641   **Vassetzky YS. 2018.** HIV-1 Tat protein induces DNA damage in human peripheral blood
642   B-lymphocytes via mitochondrial ROS production. *Redox biology* 15:97–108. DOI:
643   10.1016/j.redox.2017.11.024.

644 **Ensoli B, Barillari G, Salahuddin SZ, Gallo RC, Wong-Staal F. 1990.** Tat protein of HIV-1
645   stimulates growth of cells derived from Kaposi's sarcoma lesions of AIDS patients. *Nature*
646   345:84–86. DOI: 10.1038/345084a0.

647 **Fan Q, Liu B. 2016.** Identification of a RNA-Seq Based 8-Long Non-Coding RNA Signature
648   Predicting Survival in Esophageal Cancer. *Medical science monitor: international medical*
649   *journal of experimental and clinical research* **22:**5163–5172. DOI: 10.12659/msm.902615.

650 **Fritsch L, Marechal V, Schneider V, Barthet C, Rozenbaum W, Moisan-Coppey M,**
651   **Coppey J, Nicolas JC. 1998.** Production of HIV-1 by human B cells infected in vitro:
652   characterization of an EBV genome-negative B cell line chronically synthetizing a low level
653   of HIV-1 after infection. *Virology* 244:542–551. DOI: 10.1006/viro.1998.9120.

654 **Gautier VW, Gu L, O'Donoghue N, Pennington S, Sheehy N, Hall WW. 2009.** In vitro
655   nuclear interactome of the HIV-1 Tat protein. *Retrovirology* 6:47. DOI: 10.1186/1742-
656   4690-6-47.

657 **Germini D, Tsfasman T, Klibi M, El-Amine R, Pichugin A, Iarovaia OV, Bilhou-Nabera C,**
658   **Subra F, Bou Saada Y, Sukhanova A, Boutboul D, Raphaël M, Wiels J, Razin SV,**
659   **Bury-Moné S, Oksenhendler E, Lipinski M, Vassetzky YS. 2017.** HIV Tat induces a
660   prolonged MYC relocalization next to IGH in circulating B-cells. *Leukemia*. DOI:
661   10.1038/leu.2017.106.

662 **Gloghini A, Dolcetti R, Carbone A. 2013.** Lymphomas occurring specifically in HIV-infected
663   patients: from pathogenesis to pathology. *Seminars in cancer biology* 23:457–467. DOI:
664   10.1016/j.semcancer.2013.08.004.

665   **Gong S, Xu M, Zhang Y, Shan Y, Zhang H. 2020.** The prognostic signature and potential

666   target genes of six long non-coding RNA in laryngeal squamous cell carcinoma. *Frontiers*

667   *in genetics* **11:**413. DOI: 10.3389/fgene.2020.00413.

668   **Gorwood J, Ejlalmanesh T, Bourgeois C, Mantecon M, Rose C, Atlan M, Desjardins D, Le**

669   **Grand R, Fève B, Lambotte O, Capeau J, Béréziat V, Lagathu C. 2020.** SIV Infection

670   and the HIV Proteins Tat and Nef Induce Senescence in Adipose Tissue and Human

671   Adipose Stem Cells, Resulting in Adipocyte Dysfunction. *Cells* 9. DOI:

672   10.3390/cells9040854.

673   **He N, Liu M, Hsu J, Xue Y, Chou S, Burlingame A, Krogan NJ, Alber T, Zhou Q. 2010.**

674   HIV-1 Tat and host AFF4 recruit two transcription elongation factors into a bifunctional

675   complex for coordinated activation of HIV-1 transcription. *Molecular cell* **38:**428–438.

676   DOI: 10.1016/j.molcel.2010.04.013.

677   **Herrmann CH, Rice AP. 1993.** Specific interaction of the human immunodeficiency virus Tat

678   proteins with a cellular protein kinase. *Virology* 197:601–608. DOI:

679   10.1006/viro.1993.1634.

680   **Herrmann CH, Rice AP. 1995.** Lentivirus Tat proteins specifically associate with a cellular

681   protein kinase, TAK, that hyperphosphorylates the carboxyl-terminal domain of the large

682   subunit of RNA polymerase II: candidate for a Tat cofactor. *Journal of virology* 69:1612–

683   1620.

684   **HIV-CAUSAL Collaboration, Ray M, Logan R, Sterne JAC, Hernández-Díaz S, Robins**

685   **JM, Sabin C, Bansi L, van Sighem A, de Wolf F, Costagliola D, Lanoy E, Bucher HC,**

686   **von Wyl V, Esteve A, Casbona J, del Amo J, Moreno S, Justice A, Goulet J, Lodi S,**

687   **Phillips A, Seng R, Meyer L, Pérez-Hoyos S, García de Olalla P, Hernán MA. 2010.**

688   The effect of combined antiretroviral therapy on the overall mortality of HIV-infected

689   individuals. *AIDS* 24:123–137. DOI: 10.1097/QAD.0b013e3283324283.

690   **Hübel K. 2020.** The Changing Landscape of Lymphoma Associated with HIV Infection.

691   *Current oncology reports* 22:111. DOI: 10.1007/s11912-020-00973-0.

692   **Ivashkiv LB, Donlin LT. 2014.** Regulation of type I interferon responses. *Nature reviews.*

693   *Immunology* 14:36–49. DOI: 10.1038/nri3581.

694   **Jiang Y, Chai L, Fasae MB, Bai Y. 2018.** The role of HIV Tat protein in HIV-related

695   cardiovascular diseases. *Journal of translational medicine* 16:121. DOI: 10.1186/s12967-

696   018-1500-0.

697 **Jin C, Peng X, Xie T, Lu X, Liu F, Wu H, Yang Z, Wang J, Cheng L, Wu N. 2016.**

698   Detection of the long noncoding RNAs nuclear-enriched autosomal transcript 1 (NEAT1)

699   and metastasis associated lung adenocarcinoma transcript 1 in the peripheral blood of HIV-

700   1-infected patients. *HIV medicine* 17:68–72. DOI: 10.1111/hiv.12276.

701 **Kanehisa M, Goto S. 2000.** KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids*

702   *research* 28:27–30. DOI: 10.1093/nar/28.1.27.

703 **Kao SY, Calman AF, Luciw PA, Peterlin BM. 1987.** Anti-termination of transcription within

704   the long terminal repeat of HIV-1 by tat gene product. *Nature* **330:**489–493. DOI:

705   10.1038/330489a0

706 **Katano H, Sato Y, Hoshino S, Tachikawa N, Oka S, Morishita Y, Ishida T, Watanabe T,**

707   **Rom WN, Mori S, Sata T, Weiden MD, Hoshino Y. 2007.** Integration of HIV-1 caused

708   STAT3-associated B cell lymphoma in an AIDS patient. *Microbes and infection / Institut*

709   *Pasteur* 9:1581–1589. DOI: 10.1016/j.micinf.2007.09.008.

710 **Kim D, Langmead B, Salzberg SL. 2015.** HISAT: a fast spliced aligner with low memory

711   requirements. *Nature methods* 12:357–360. DOI: 10.1038/nmeth.3317.

712 **Kundu RK, Sangiorgi F, Wu LY, Pattengale PK, Hinton DR, Gill PS, Maxson R. 1999.**

713   Expression of the human immunodeficiency virus-Tat gene in lymphoid tissues of

714   transgenic mice is associated with B-cell lymphoma. *Blood* 94:275–282.

715 **Krackhardt AM, Witzens M, Harig S, Hodi FS, Zauls AJ, Chessia M, Barrett P, Gribben**

716   **JG. 2002.** Identification of tumor-associated antigens in chronic lymphocytic leukemia by

717   SEREX. *Blood* **100:**2123–2131. DOI: 10.1182/blood-2002-02-0513.

718 **Kurnaeva MA, Zalevsky AO, Arifulin EA, Lisitsyna OM, Tvorogova AV, Shubina MY,**

719   **Bourenkov GP, Tikhomirova MA, Potashnikova DM, Kachalova AI, Musinova YR,**

720   **Golovin AV, Vassetzky YS, Sheval EV. 2022.** Molecular coevolution of nuclear and

721   nucleolar localization signals inside the basic domain of HIV-1 Tat. *Journal of virology*

722   96:e0150521. DOI: 10.1128/JVI.01505-21.

723 **Lazzi S, Bellan C, De Falco G, Cinti C, Ferrari F, Nyongo A, Claudio PP, Tosi GM, Vatti**

724   **R, Gloghini A, Carbone A, Giordano A, Leoncini L, Tosi P. 2002.** Expression of

725   RB2/p130 tumor-suppressor gene in AIDS-related non-Hodgkin's lymphomas: implications

726   for disease pathogenesis. *Human pathology* 33:723–731. DOI: 10.1053/hupa.2002.125372.

727   **Li H, An X, Li Q, Yu H, Li Z. 2021.** Construction and analysis of competing endogenous RNA
728   network of MCF-7 breast cancer cells based on the inhibitory effect of 6-thioguanine on cell
729   proliferation. *Oncology letters* **21:**104. DOI: 10.3892/ol.2020.12365.

730   **Li H, Gao J, Liu L, Zhang S. 2022.** LINC00958: A promising long non-coding RNA related to
731   cancer. *Biomedicine & pharmacotherapy* **151:**113087. DOI: 10.1016/j.biopha.2022.113087.

732   **Love MI, Huber W, Anders S. 2014.** Moderated estimation of fold change and dispersion for
733   RNA-seq data with DESeq2. *Genome biology* 15:550. DOI: 10.1186/s13059-014-0550-8.

734   **Marban C, Su T, Ferrari R, Li B, Vatakis D, Pellegrini M, Zack JA, Rohr O, Kurdistani**
735   **SK. 2011.** Genome-wide binding map of the HIV-1 Tat protein to the human genome. *PloS*
736   *one* 6:e26894. DOI: 10.1371/journal.pone.0026894.

737   **Marino J, Maubert ME, Mele AR, Spector C, Wigdahl B, Nonnemacher MR. 2020.**
738   Functional impact of HIV-1 Tat on cells of the CNS and its role in HAND. *Cellular and*
739   *molecular life sciences: CMLS.* DOI: 10.1007/s00018-020-03561-4.

740   **Mateen FJ, Shinohara RT, Carone M, Miller EN, McArthur JC, Jacobson LP, Sacktor N,**
741   **Multicenter AIDS Cohort Study (MACS) Investigators. 2012.** Neurologic disorders
742   incidence in HIV+ vs HIV- men: Multicenter AIDS Cohort Study, 1996-2011. *Neurology*
743   79:1873–1880. DOI: 10.1212/WNL.0b013e318271f7b8.

744   **McCune JM, Fu SM, Kunkel HG. 1981.** J chain biosynthesis in pre-B cells and other possible
745   precursor B cells. *The Journal of experimental medicine* 154:138–145. DOI:
746   10.1084/jem.154.1.138.

747   **Mohammed A, Biegert G, Adamec J, Helikar T. 2017.** Identification of potential tissue-
748   specific cancer biomarkers and development of cancer versus normal genomic classifiers.
749   *Oncotarget* **8:**85692–85715. DOI: 10.18632/oncotarget.21127.

750   **Muniz L, Egloff S, Ughy B, Jády BE, Kiss T (2010)** Controlling Cellular P-TEFb Activity by
751   the HIV-1 Transcriptional Transactivator Tat. PLoS Pathog **6(10):** e1001152.
752   https://doi.org/10.1371/journal.ppat.1001152.

753   **Musinova YR, Sheval EV, Dib C, Germini D, Vassetzky YS. 2016.** Functional roles of HIV-1
754   Tat protein in the nucleus. *Cellular and molecular life sciences: CMLS* 73:589–601. DOI:
755   10.1007/s00018-015-2077-x.

756   **Nath A. 2015.** Eradication of human immunodeficiency virus from brain reservoirs. *Journal of*
757   *neurovirology* 21:227–234. DOI: 10.1007/s13365-014-0291-1.

758 **Nekhai S, Jeang K-T. 2006.** Transcriptional and post-transcriptional regulation of HIV-1 gene
759   expression: role of cellular factors for Tat and Rev. *Future microbiology* **1:**417–426. DOI:
760   10.2217/17460913.1.4.417.

761 **Nookala AR, Kumar A. 2014.** Molecular mechanisms involved in HIV-1 Tat-mediated
762   induction of IL-6 and IL-8 in astrocytes. *Journal of neuroinflammation* 11:214. DOI:
763   10.1186/s12974-014-0214-3.

764 **Noy A. 2020.** HIV Lymphoma and Burkitts Lymphoma. *Cancer journal* 26:260–268. DOI:
765   10.1097/PPO.0000000000000448.

766 **Petri A, Dybkær K, Bøgsted M, Thrue CA, Hagedorn PH, Schmitz A, Bødker JS, Johnsen**
767   **HE, Kauppinen S. 2015.** Long Noncoding RNA Expression during Human B-Cell
768   Development. *PloS one* **10:**e0138236. DOI: 10.1371/journal.pone.0138236.

769 **Pham VV, Salguero C, Khan SN, Meagher JL, Brown WC, Humbert N, de Rocquigny H,**
770   **Smith JL, D'Souza VM. 2018.** HIV-1 Tat interactions with cellular 7SK and viral TAR
771   RNAs identifies dual structural mimicry. *Nature communications* **9:**4266. DOI:
772   10.1038/s41467-018-06591-6.

773 **Poggi A, Carosio R, Fenoglio D, Brenci S, Murdaca G, Setti M, Indiveri F, Scabini S,**
774   **Ferrero E, Zocchi MR. 2004.** Migration of V delta 1 and V delta 2 T cells in response to
775   CXCR3 and CXCR4 ligands in healthy donors and HIV-1-infected patients: competition by
776   HIV-1 Tat. *Blood* 103:2205–2213. DOI: 10.1182/blood-2003-08-2928.

777 **Qu D, Sun W-W, Li L, Ma L, Sun L, Jin X, Li T, Hou W, Wang J-H. 2019.** Long noncoding
778   RNA MALAT1 releases epigenetic silencing of HIV-1 replication by displacing the
779   polycomb repressive complex 2 from binding to the LTR promoter. *Nucleic acids research*
780   47:3013–3027. DOI: 10.1093/nar/gkz117.

781 **Pushkarsky T, Ward A, Ivanov A, Lin X, Sviridov D, Nekhai S, Bukrinsky MI. 2022.**
782   Abundance of Nef and p-Tau217 in Brains of Individuals Diagnosed with HIV-Associated
783   Neurocognitive Disorders Correlate with Disease Severance. *Molecular neurobiology*
784   **59:**1088–1097. DOI: 10.1007/s12035-021-02608-2.

785 **Raymond AD, Campbell-Sims TC, Khan M, Lang M, Huang MB, Bond VC, Powell MD.**
786   **2011.** HIV Type 1 Nef is released from infected cells in CD45(+) microvesicles and is
787   present in the plasma of HIV-infected individuals. *AIDS research and human retroviruses*
788   **27:**167–178. DOI: 10.1089/aid.2009.0170.

789 **Reeder JE, Kwak Y-T, McNamara RP, Forst CV, D'Orso I. 2015.** HIV Tat controls RNA
790   Polymerase II and the epigenetic landscape to transcriptionally reprogram target immune
791   cells. *eLife* 4. DOI: 10.7554/eLife.08955.

792 **Rhim H, Echetebu CO, Herrmann CH, Rice AP. 1994.** Wild-type and mutant HIV-1 and
793   HIV-2 Tat proteins expressed in Escherichia coli as fusions with glutathione S-transferase.
794   *Journal of acquired immune deficiency syndromes* 7:1116–1121.

795 **Sall FB, El Amine R, Markozashvili D, Tsfasman T, Oksenhendler E, Lipinski M,**
796   **Vassetzky Y, Germini D. 2019.** HIV-1 Tat protein induces aberrant activation of AICDA
797   in human B-lymphocytes from peripheral blood. *Journal of cellular physiology*. DOI:
798   10.1002/jcp.28219.

799 **Sedore SC, Byers SA, Biglione S, Price JP, Maury WJ, Price DH. 2007.** Manipulation of P-
800   TEFb control machinery by HIV: recruitment of P-TEFb from the large form by Tat and
801   binding of HEXIM1 to TAR. *Nucleic acids research* **35:**4347–4358. DOI:
802   10.1093/nar/gkm443.

803 **Shmakova A, Germini D, Vassetzky Y. 2020.** HIV-1, HAART and cancer: A complex
804   relationship. *International journal of cancer. Journal international du cancer* 146:2666–
805   2679. DOI: 10.1002/ijc.32730.

806 **Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich**
807   **A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. 2005.** Gene set enrichment analysis:
808   a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings*
809   *of the National Academy of Sciences of the United States of America* 102:15545–15550.
810   DOI: 10.1073/pnas.0506580102.

811 **Supek F, Bošnjak M, Škunca N, Šmuc T. 2011.** REVIGO summarizes and visualizes long lists
812   of gene ontology terms. *PloS one* 6:e21800. DOI: 10.1371/journal.pone.0021800.

813 **The Gene Ontology Consortium. 2017.** Expansion of the Gene Ontology knowledgebase and
814   resources. *Nucleic acids research* 45:D331–D338. DOI: 10.1093/nar/gkw1108.

815 **Vogel J, Hinrichs SH, Napolitano LA, Ngo L, Jay G. 1991.** Liver cancer in transgenic mice
816   carrying the human immunodeficiency virus tat gene. *Cancer research* 51:6686–6690.

817 **Vogel J, Hinrichs SH, Reynolds RK, Luciw PA, Jay G. 1988.** The HIV tat gene induces
818   dermal lesions resembling Kaposi's sarcoma in transgenic mice. *Nature* 335:606–611. DOI:
819   10.1038/335606a0.

820 **Wang W, Guo M, Hu L, Cai J, Zeng Y, Luo J, Shu Z, Li W, Huang Z. 2012.** The zinc finger
821   protein ZNF268 is overexpressed in human cervical cancer and contributes to tumorigenesis
822   via enhancing NF-κB signaling. *The Journal of biological chemistry* **287:**42856–42866.
823   DOI: 10.1074/jbc.M112.399923.

824 **Wang T, Yi R, Green LA, Chelvanambi S, Seimetz M, Clauss M. 2015.** Increased
825   cardiovascular disease risk in the HIV-positive population on ART: potential role of HIV-
826   Nef and Tat. *Cardiovascular pathology: the official journal of the Society for*
827   *Cardiovascular Pathology* 24:279–282. DOI: 10.1016/j.carpath.2015.07.001.

828 **Wang G, Zheng C. 2021.** Zinc finger proteins in the host-virus interplay: multifaceted functions
829   based on their nucleic acid-binding property. *FEMS microbiology reviews* 45. DOI:
830   10.1093/femsre/fuaa059.

831 **West JA, Davis CP, Sunwoo H, Simon MD, Sadreyev RI, Wang PI, Tolstorukov MY,**
832   **Kingston RE. 2014.** The long noncoding RNAs NEAT1 and MALAT1 bind active
833   chromatin sites. *Molecular cell* 55:791–802. DOI: 10.1016/j.molcel.2014.07.012.

834 **Westendorp MO, Frank R, Ochsenbauer C, Stricker K, Dhein J, Walczak H, Debatin KM,**
835   **Krammer PH. 1995.** Sensitization of T cells to CD95-mediated apoptosis by HIV-1 Tat
836   and gp120. *Nature* 375:497–500. DOI: 10.1038/375497a0.

837 **Xiao H, Neuveut C, Tiffany HL, Benkirane M, Rich EA, Murphy PM, Jeang KT. 2000.**
838   Selective CXCR4 antagonism by Tat: implications for in vivo expansion of coreceptor use
839   by HIV-1. *Proceedings of the National Academy of Sciences of the United States of*
840   *America* 97:11466–11471. DOI: 10.1073/pnas.97.21.11466.

841 **Yu G, Wang L-G, Han Y, He Q-Y. 2012.** clusterProfiler: an R package for comparing
842   biological themes among gene clusters. *Omics: a journal of integrative biology* 16:284–287.
843   DOI: 10.1089/omi.2011.0118.

844 **Yue H, Wu K, Liu K, Gou L, Huang A, Tang H. 2022.** LINC02154 promotes the proliferation
845   and metastasis of hepatocellular carcinoma by enhancing SPC24 promoter activity and
846   activating the PI3K-AKT signaling pathway. *Cellular oncology* 45:447–462. DOI:
847   10.1007/s13402-022-00676-7.

848 **Zhang Q, Chen C-Y, Yedavalli VSR, Jeang K-T. 2013.** NEAT1 long noncoding RNA and
849   paraspeckle bodies modulate HIV-1 posttranscriptional expression. *mBio* 4. DOI:
850   10.1128/mbio.00596-12.

851   **Zhang Y, Jin T, Shen H, Yan J, Guan M, Jin X. 2019.** Identification of Long Non-Coding

852   RNA Expression Profiles and Co-Expression Genes in Thyroid Carcinoma Based on The

853   Cancer Genome Atlas (TCGA) Database. *Medical science monitor: international medical*

854   *journal of experimental and clinical research* **25:**9752–9769. DOI: 10.12659/MSM.917845.

855   **Zhao Z, Wang D, Zhu C, Shao H, Sun C, Qiu H, Xue L, Xu J, Guo M, Li W. 2008.**

856   Aberrant alternative splicing of human zinc finger gene ZNF268 in human hematological

857   malignancy. *Oncology reports* **20:**1243–1248.

858   **Zhong F, Zhu M, Gao K, Xu P, Yang H, Hu D, Cui D, Wang M, Xie X, Wei Y, Zhang H,**

859   **Du H. 2018.** Low expression of the long non-coding RNA NR_026827 in gastric cancer.

860   *American journal of translational research* **10:**2706–2711.

861   **Zuo S, Wang L, Wen Y, Dai G. 2018.** Identification of a universal 6-lncRNA prognostic

862   signature for three pathologic subtypes of renal cell carcinoma. *Journal of cellular*

863   *biochemistry.* **120:**7375-7385. DOI: 10.1002/jcb.28012.

864 **Figure legends**

865 **Figure 1**

866 **Cell lines that were used to analyze HIV-1 Tat action of cultured B cells.**

867 (A) Four cell lines that were used in this study. Created with BioRender.com. (B) EGFP

868 fluorescence of demonstrated high purity and homogeneity of the obtained cell lines (cells

869 without EGFP fluorescence are colored blue, and those with EGFP fluorescence are colored

870 green). (C) Western blot analysis of EGFP, Tat-EGFP and TatC22G-EGFP expression in the cell

871 lines. (D) The transactivation ability of Tat-EGFP in HeLa cells with integrated LTR-TurboRFP

872 (flow cytometry, a representative experiment). TurboRFP fluorescence was clearly detected after

873 the expression of Tat-EGFP but not in nontransduced cells (control) or after the expression of

874 EGFP.

875

876 **Figure 2**

877 **Differentially expressed genes (DEGs) in RPMI cells expressing EGFP, EGFP-Tat or**

878 **EGFP-TatC22G.**

879 (A) The number of all DEGs (left) and protein-coding DEGs (right) found in three comparisons:

880 $RPMI^{EGFP}$ versus RPMI, $RPMI^{Tat}$ versus RPMI, and $RPMI^{Cys}$ versus RPMI. (B) Validation of the

881 RNA-seq dataset using qRT-PCR on the indicated upregulated and downregulated genes (mean

882 ± SEM; n = 3). (C) Upregulated (left) and downregulated (right) GO BP terms affected by Tat

883 identified by GSEA ($RFPI^{Tat}$ cells versus RPMI cells). Only significantly enriched (adjusted p

884 value < 0.05) and nonredundant GO BP terms are shown (the top 20).

885

886 **Figure 3**

887 **Enrichment analysis of Tat-affected DEGs ($RPMI^{Tat}$ versus RPMI).**

888 (A) KEGG pathways positively regulated by Tat, identified by ORA of protein-coding DEGs.

889 Only significantly enriched (adjusted p value < 0.05) KEGG pathways are shown. (B) Activated

890 KEGG pathways (hsa04620, hsa04380, hsa04621, and hsa05168) and associated DEGs after

891 filtering overlapping gene sets. (C) KEGG pathways negatively regulated by Tat, identified by

892 overrepresentation analysis of protein-coding DEGs. Only significantly enriched (adjusted p

893 value < 0.05) KEGG pathways are shown. (D) Suppressed KEGG pathways (hsa04514,

894 hsa04060, and hsa04670) and associated DEGs after filtering overlapping gene sets.

895

896   **Figure 4**

897   **Comparison of gene expression between RPMI^{Tat} and RPMI^{Cys} cells.**

898   (A) Venn diagrams demonstrated that approximately half of the DEGs overlapped, indicating

899   similar but not identical modifications induced by EGFP-Tat and its mutated form EGFP-

900   TatC22G. (B) Suppressed KEGG pathways (hsa04060, hsa04080, hsa05033, and hsa05032) and

901   associated DEGs when comparing RPMI^{Cys} against RPMI^{Tat} cells.

902

903   **Figure 5**

904   **EGFP, Tat-EGFP or TatC22G-EGFP expression does not affect the morphology or**

905   **proliferative potential of RPMI 8866 cells.**

906   (A) Representative cells of RPMI, RPMI^{EGFP}, RPMI^{Tat}, and RPMI^{Cys} lines under electron

907   microscopy. Bars = 1 μm. (B) Cell cycle distribution of RPMI, RPMI^{EGFP}, RPMI^{Tat}, and RPMI^{Cys}

908   cells (representative experiment) and (C) estimation of cell proportions at the G0/G1, S, and

909   G2/M stages (mean ± SD, n = 3). (D) Estimation of S-phase cells using incorporation of

910   synthetic nucleotides (EdU) (mean ± SD, n = 3). (E) Estimation of cycling cells (Ki-67-positive)

911   (mean ± SD, n = 3). (F) Estimation of the content of apoptotic cells in RPMI, RPMI^{EGFP},

912   RPMI^{Tat}, and RPMI^{Cys} lines using four independent methods (mean ± SD, n = 3). In Panels C, D,

913   E, and F, all differences between control cells (RPMI) and cells expressing different proteins

914   (RPMI^{EGFP}, RPMI^{Tat}, and RPMI^{Cys}) were insignificant (Kruskal–Wallis test, p > 0.05; n = 3).

915

916   **Figure 6**

917   **HIV-1 Tat influences cellular dynamics and chromosome organization upon prolonged**

918   **cultivation.**

919   (A) Phase contrast images and EGFP fluorescence (merged) in RPMI, RPMI^{EGFP}, RPMI^{Tat}, and

920   RPMI^{Cys} after 3 months of cultivation (representative images). In the RPMI^{Tat} line, there are

921   ~40% cells without EGFP fluorescence (arrowheads). (B) The proportion of EGFP-positive cells

922   was decreased during prolonged cultivation of RPMI^{Tat} but not RPMI^{EGFP} and RPMI^{Cys} cells (a

923   representative experiment). (C) Representative image of a metaphase plate with detected

924   chromosomes 1 (red), 2 (green) and 4 (orange). The chromosome with a translocation is marked

925   with an arrow. (D) Cells expressing Tat-EGFP or TatC22G-EGFP contained a significantly

926 higher proportion of chromosomes with aberrations (see Table S6 for a detailed description of
927 chromosome aberrations). The comparison was performed using Fisher's exact test. Differences
928 were considered statistically significant (*) at p <0.01.

Upregulated and downregulated long non-coding RNAs (PPMI^{Tat} vs RPMI cells) with the highest expression (top 10). The full list of DE lncRNAs is presented in Table S4.

1   **Table 1**

2   **Upregulated and downregulated long non-coding RNAs (PPMI^{Tat} vs RPMI cells) with the**

3   **highest expression (top 10). The full list of DE lncRNAs is presented in Table S4.**

Upregulated genes							
EnsID	Gene name	padj	log2FC	FC	Mean count	Functions and possible involvement in oncogenesis	Reference
ENSG00000251562	MALAT1	9.9E-22	1.28	2.44	34 783.28	Supplementary Table S5	
ENSG00000245532	NEAT1	2.2E-08	0.73	1.66	3 039.09	Supplementary Table S5	
ENSG00000260658	RP11-368L12.1	7.0E-05	0.67	1.59	1 545.91	Co-expressed with gene module of actively proliferating pre-B cells	(Petri et al., 2015)
ENSG00000248837	RP11-412P11.1	9.5E-08	0.99	1.99	859.73	-	
ENSG00000230448	LINC00276	2.6E-04	0.59	1.50	517.66	-	
ENSG00000234663	LINC01934	1.4E-09	0.90	1.86	516.91	Downregulated in thyroid carcinoma	(Zhang et al., 2019)
ENSG00000226965	AC003088.1	2.3E-11	1.19	2.29	319.39	Upregulated in breast cancer line	(Li et al., 2021)

ENSG000002381 29	RP3-410C9.2	8.2E-15	0.86	1.82	307.45	-	
ENSG000002353 85	LINC02154	8.5E-79	2.43	5.40	272.56	Upregulated in Laryngeal Squamous Cell Carcinoma and renal cell carcinoma, and can be used as a prognostic feature; promotes the proliferation and metastasis of hepatocellular carcinoma	(Zuo et al., 2018; Gong et al., 2020; Yue et al., 2022)
ENSG000002496 45	RP11-552M14.1	2.0E-22	1.19	2.28	261.05	-	
Downregulated genes							
ENSG000002533 77	RP11-566H8.3	6.4E-07	-0.86	0.55	1 751.51	Produces non-canonical cancer/testis antigen	(Chong et al., 2020)
ENSG000002541 19	RP11-705O24.1	9.9E-16	-1.40	0.38	1 623.69	Potential prognostic feature in esophageal cancer	(Fan & Liu, 2016)
ENSG000002513 81	LINC00958	9.4E-23	-0.68	0.62	1 003.09	Canonical lncRNA in human cancer progression (overexpressed in many cancers)	(Li et al., 2022)
ENSG000002677	CTD-	4.5E-08	-1.13	0.46	895.69	Potential cancer-	(Mohammed et

61	2130O13.1					specific biomarker	al., 2017)
ENSG000002719 55	RP11-444A22.1	3.4E-79	-3.03	0.12	803.23	Involved in acute myeloid leukemia cell differentiation	(Cozzi et al., 2022)
ENSG000002317 72	RP1-154K9.2	8.3E-53	-9.37	0.00	726.17	-	
ENSG000002276 81	RP11-307P5.1	4.5E-10	-1.17	0.44	621.53	-	
ENSG000002510 88	RP11-325B23.2	1.5E-04	-0.63	0.65	454.27	-	
ENSG000001859 04	LINC00839	2.5E-62	-1.88	0.27	332.95	Downregulated in gastric cancer with *H. pilori* infection; upregulated in chemoresistant breast cancer	(Zhong et al., 2018; Chen et al., 2020)
ENSG000002427 41	LINC02005	1.6E-20	-1.44	0.37	324.00	-	

Figure 1

Cell lines that were used to analyze HIV-1 Tat action of cultured B cells

(A) Four cell lines that were used in this study. Created with BioRender.com. (B) EGFP
fluorescence of demonstrated high purity and homogeneity of the obtained cell lines (cells
without EGFP fluorescence are colored blue, and those with EGFP fluorescence are colored
green). (C) Western blot analysis of EGFP, Tat-EGFP and TatC22G-EGFP expression in the cell
lines. (D) The transactivation ability of Tat-EGFP in HeLa cells with integrated LTR-TurboRFP
(flow cytometry, a representative experiment). TurboRFP fluorescence was clearly detected
after the expression of Tat-EGFP but not in nontransduced cells (control) or after the
expression of EGFP.
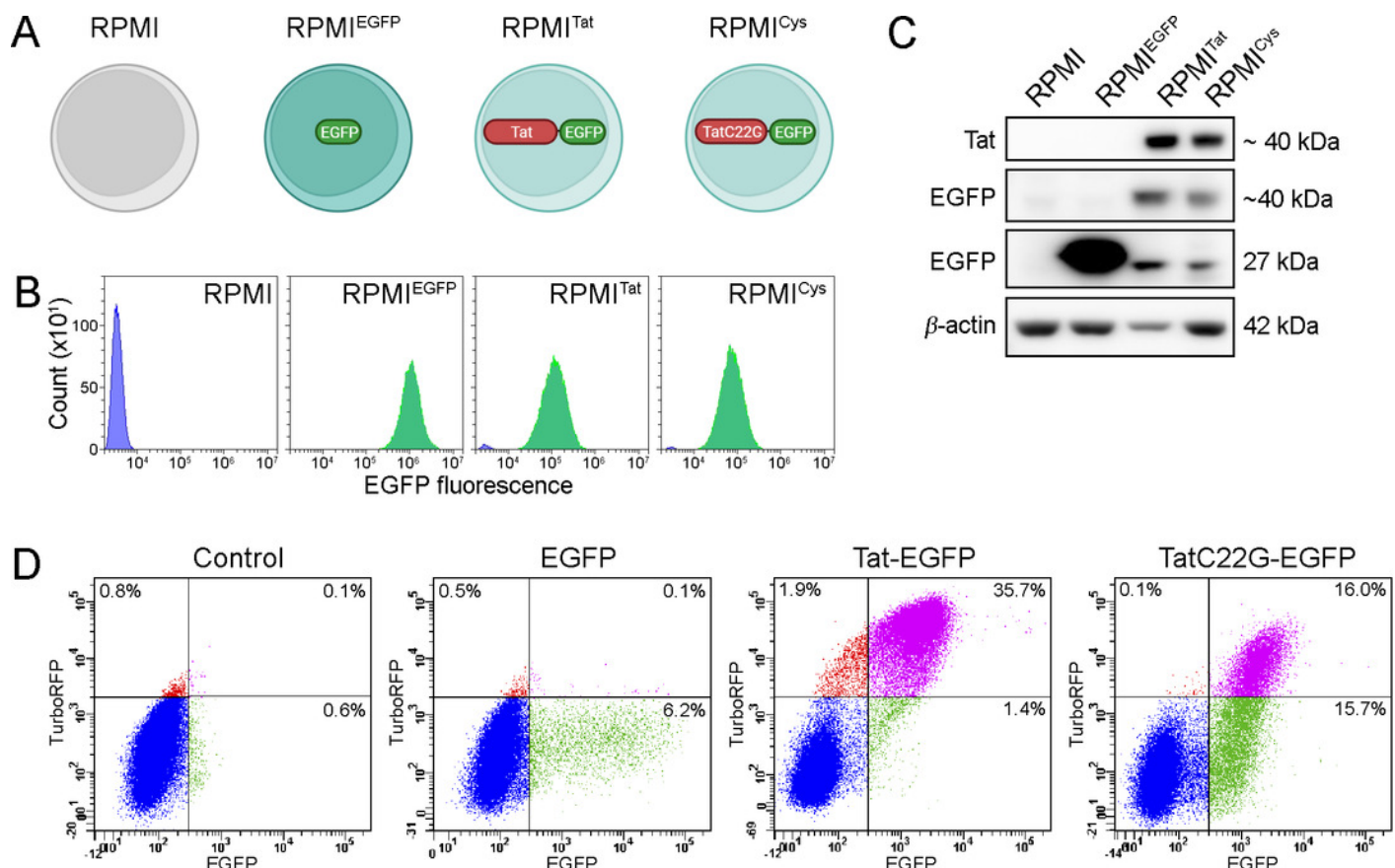
Figure 2

Differentially expressed genes (DEGs) in RPMI cells expressing EGFP, EGFP-Tat or EGFP-TatC22G.

(A) The number of all DEGs (left) and protein-coding DEGs (right) found in three comparisons: RPMI^{EGFP} versus RPMI, RPMI^{Tat} versus RPMI, and RPMI^{Cys} versus RPMI. (B) Validation of the RNA-seq dataset using qRT-PCR on the indicated upregulated and downregulated genes (mean ± SEM; n = 3). (C) Upregulated (left) and downregulated (right) GO BP terms affected by Tat identified by GSEA (RFPI^{Tat} cells versus RPMI cells). Only significantly enriched (adjusted p value < 0.05) and nonredundant GO BP terms are shown (the top 20).
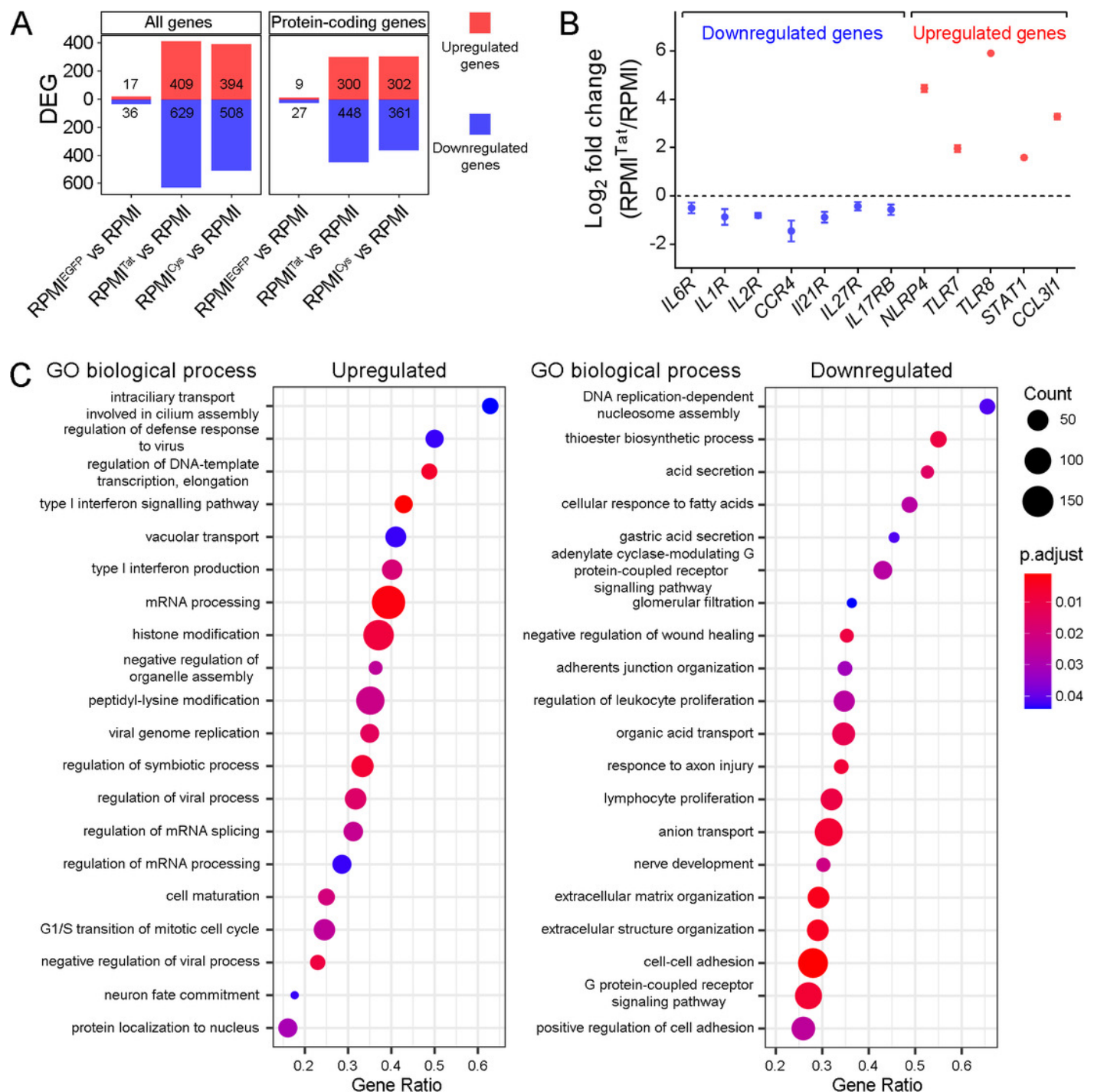
Figure 3

Enrichment analysis of Tat-affected DEGs (RPMI^{Tat} versus RPMI)

(A) KEGG pathways positively regulated by Tat, identified by ORA of protein-coding DEGs. Only significantly enriched (adjusted p value < 0.05) KEGG pathways are shown. (B) Activated KEGG pathways (hsa04620, hsa04380, hsa04621, and hsa05168) and associated DEGs after filtering overlapping gene sets. (C) KEGG pathways negatively regulated by Tat, identified by overrepresentation analysis of protein-coding DEGs. Only significantly enriched (adjusted p value < 0.05) KEGG pathways are shown. (D) Suppressed KEGG pathways (hsa04514, hsa04060, and hsa04670) and associated DEGs after filtering overlapping gene sets.
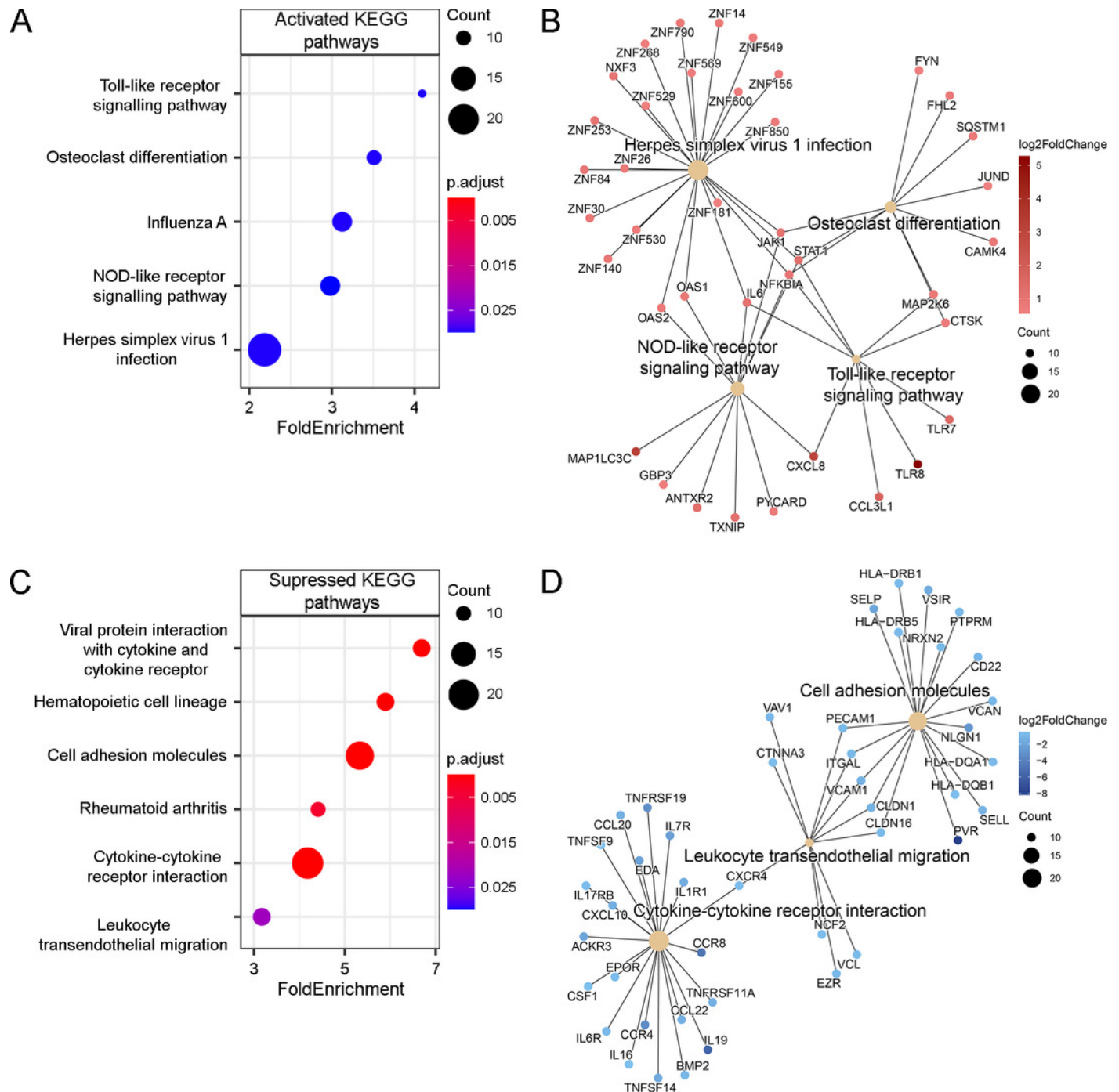
Figure 4

Comparison of gene expression between RPMI^{Tat} and RPMI^{Cys} cells

(A) Venn diagrams demonstrated that approximately half of the DEGs overlapped, indicating similar but not identical modifications induced by EGFP-Tat and its mutated form EGFP-TatC22G. (B) Suppressed KEGG pathways (hsa04060, hsa04080, hsa05033, and hsa05032) and associated DEGs when comparing RPMI^{Cys} against RPMI^{Tat} cells.
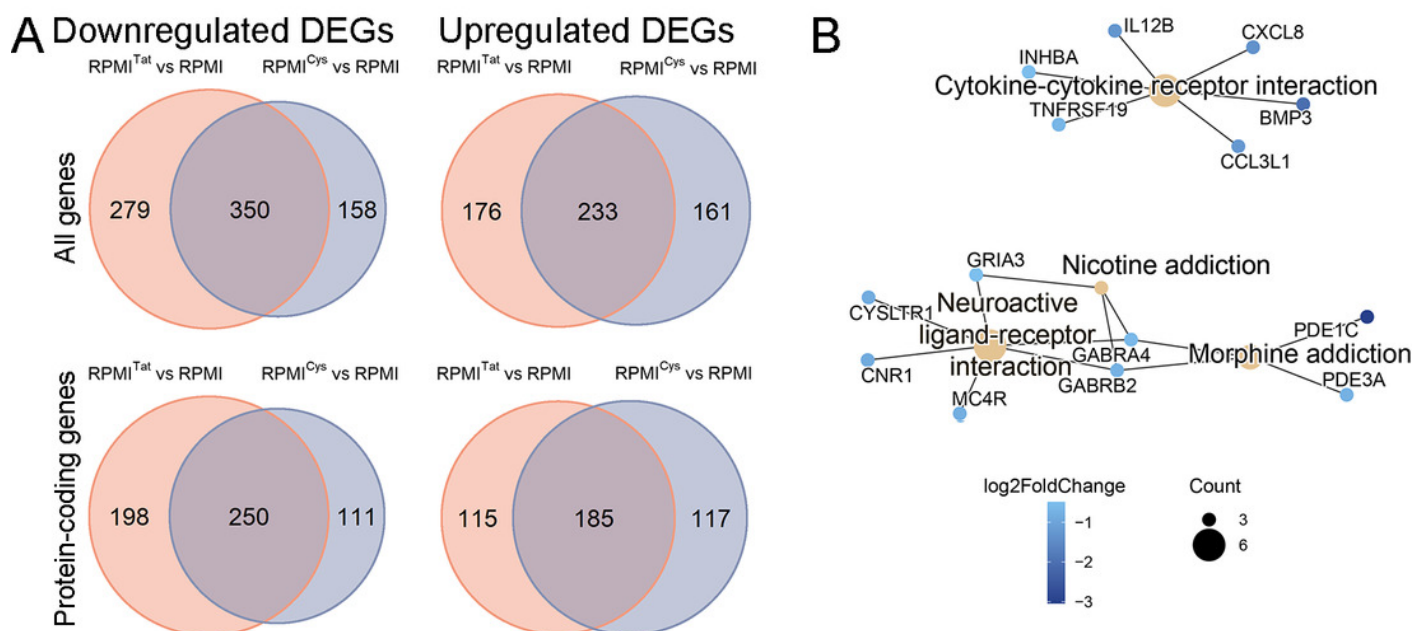
Figure 5

EGFP, Tat-EGFP or TatC22G-EGFP expression does not affect the morphology or proliferative potential of RPMI 8866 cells

(A) Representative cells of RPMI, RPMI^{EGFP}, RPMI^{Tat}, and RPMI^{Cys} lines under electron

microscopy. Bars = 1 µm. (B) Cell cycle distribution of RPMI, RPMI^{EGFP}, RPMI^{Tat}, and RPMI^{Cys}

cells (representative experiment) and (C) estimation of cell proportions at the G0/G1, S, and

G2/M stages (mean ± SD, n = 3). (D) Estimation of S-phase cells using incorporation of

synthetic nucleotides (EdU) (mean ± SD, n = 3). (E) Estimation of cycling cells (Ki-67-

positive) (mean ± SD, n = 3). (F) Estimation of the content of apoptotic cells in RPMI,

RPMI^{EGFP}, RPMI^{Tat}, and RPMI^{Cys} lines using four independent methods (mean ± SD, n = 3). In

Panels C, D, E, and F, all differences between control cells (RPMI) and cells expressing

different proteins (RPMI^{EGFP}, RPMI^{Tat}, and RPMI^{Cys}) were insignificant (Kruskal–Wallis test, p >

0.05; n = 3).
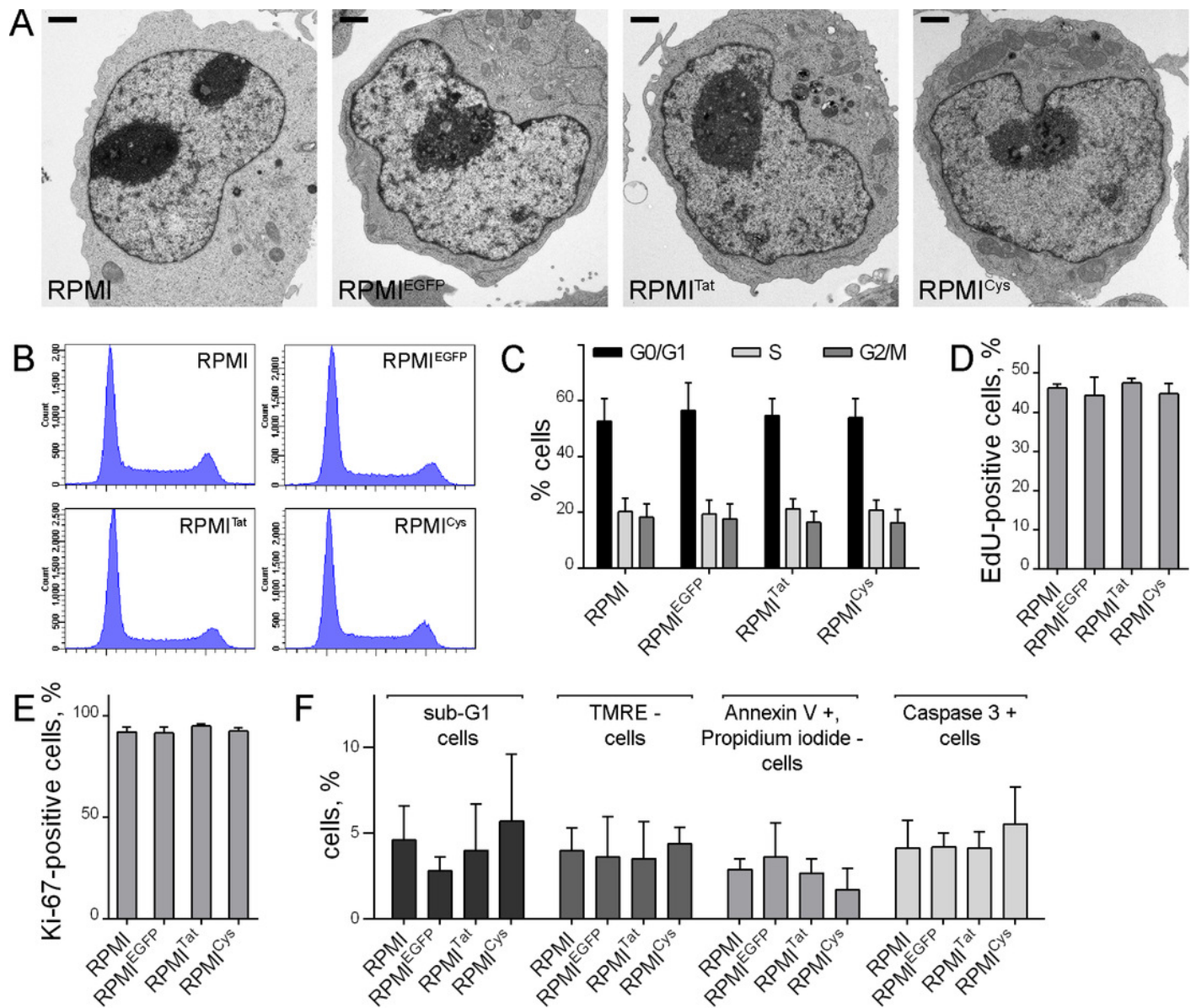
Figure 6

HIV-1 Tat influences cellular dynamics and chromosome organization upon prolonged
cultivation

(A) Phase contrast images and EGFP fluorescence (merged) in RPMI, RPMI^{EGFP}, RPMI^{Tat}, and

RPMI^{Cys} after 3 months of cultivation (representative images). In the RPMI^{Tat} line, there are

~40% cells without EGFP fluorescence (arrowheads). (B) The proportion of EGFP-positive cells

was decreased during prolonged cultivation of RPMI^{Tat} but not RPMI^{EGFP} and RPMI^{Cys} cells (a

representative experiment). (C) Representative image of a metaphase plate with detected

chromosomes 1 (red), 2 (green) and 4 (orange). The chromosome with a translocation is

marked with an arrow. (D) Cells expressing Tat-EGFP or TatC22G-EGFP contained a

significantly higher proportion of chromosomes with aberrations (see Table S6 for a detailed

description of chromosome aberrations). The comparison was performed using Fisher's exact

test. Differences were considered statistically significant (*) at $p < 0.01$.