Sporadic on/off switching of HTLV-1 Tax expression is crucial to maintain the whole population of virus-induced leukemic cells

（HTLV-1 Tax の散発的な一過性発現はウイルス性白血病細胞の集団の維持に必要である）

Mohamed Mahgoub Mohamed Ahmed Mohamed

モハメド マジュブ モハメド アーメド モハメド

Proc. Natl. Acad. Sci. U. S. A.

Sporadic on/off switching of HTLV-1 Tax expression is crucial to maintain the whole population of virus-induced leukemic cells

Mohamed Mahgoub[a], Jun-ichirou Yasunaga[a,1], Shingo Iwami[b,c,d], Shinji Nakaoka[c,e], Yoshiki Koizumi[f], Kazuya Shimura[a], and Masao Matsuoka[a,g,1]

[a]Laboratory of Virus Control, Institute for Frontier Life and Medical Sciences, Kyoto University, Kyoto, Japan; [b]Mathematical Biology Laboratory, Department of Biology, Faculty of Sciences, Kyushu University, Fukuoka, Japan; [c]PRESTO, Japan Science and Technology Agency, Kawaguchi, Saitama, Japan; [d]CREST, Japan Science and Technology Agency, Kawaguchi, Saitama, Japan; [e]Institute of Industrial Sciences, The University of Tokyo, Tokyo, Japan; [f]School of Medicine, College of Medical, Pharmaceutical and Health Sciences, Kanazawa University, Ishikawa, Japan; [g]Department of Hematology, Rheumatology and Infectious Diseases, Kumamoto University School of Medicine, Kumamoto, Japan.

[1]To whom correspondence should be addressed

Jun-ichirou Yasunaga

Tel: 81-75-751-4048; Fax: 81-75-751-4049; e-mail: jyasunag@infront.kyoto-u.ac.jp

Masao Matsuoka

Tel: 81-96-373-5156; Fax: 81-96-373-515 8 ; e-mail: mamatsu@kumamoto-u.ac.jp

Key words: HTLV-1, Tax, HBZ, adult T-cell leukemia-lymphoma, computational simulation

Abstract

Viruses causing chronic infection artfully manipulate infected cells to enable viral persistence *in vivo* under the pressure of immunity. Human T-cell leukemia virus type 1 (HTLV-1) establishes persistent infection mainly in $CD4^+$ T cells *in vivo* and induces leukemia in this subset. HTLV-1-encoded Tax is a critical transactivator of viral replication and a potent oncoprotein, but its significance in pathogenesis remains obscure due to its very low level of expression *in vivo*. Here, we show that Tax is expressed in a minor fraction of leukemic cells at any given time, and importantly, its expression spontaneously switches between on and off states. Live-cell imaging revealed that the average duration of one episode of Tax expression is approximately 19 hours. Knockdown of Tax rapidly induced apoptosis in most cells, indicating that Tax is critical for maintaining the population even if its short-term expression is limited to a small sub-population. Single-cell analysis and computational simulation suggest that transient Tax expression triggers anti-apoptotic machinery and this effect continues even after Tax expression is diminished; this activation of the anti-apoptotic machinery is the critical event for maintaining the population. In addition, Tax is induced by various cytotoxic stresses and also promotes HTLV-1 replication. Thus it appears that Tax protects infected cells from apoptosis and increases the chance of viral transmission at a critical moment. Keeping the expression of Tax minimal but inducible on demand is therefore a fundamental strategy of HTLV-1 to promote persistent infection and leukemogenesis.

Significance Statement

The oncogenic retrovirus HTLV-1 encodes Tax, an activator of both viral replication and cellular oncogenic pathways. Despite the potent activities of Tax, its precise roles in pathogenesis remains unclear, since it is faintly expressed *in vivo*. This study shows that sporadic and transient Tax expression is observed in a small subpopulation of HTLV-1-induced leukemic cells. This limited Tax expression is critical for survival of the whole population through ignition of anti-apoptotic signals. Tax is induced by various stresses, suggesting that Tax efficiently protect cells from cell death and reactivates the virus from reservoirs under condition of cellular stress. It is an elaborated strategy of HTLV-1 to evade host immunity and enable persistence *in vivo*.

/body

Chronic viral infection is established when there is a metastable balance between host immunity and viral strategies for maintaining infected cells in infected individuals. Several human viruses cause persistent infection and are closely associated with inflammatory diseases and/or cancers (1). Human T-cell leukemia virus type 1 (HTLV-1) and human herpesviruses, such as Epstein-Barr virus (EBV/HHV-4) and Kaposi's sarcoma-associated virus (KSHV/HHV-8), are representatives of this type of viruses (2), and express a limited number/amount of viral products, enabling infected cells to survive *in vivo* by escaping from host immune surveillance. Reactivation of viral replication, such as in the lytic phase of EBV and KSHV infection, is then a critical step for facilitating viral transmission to new hosts. Human oncogenic viruses encode the regulatory factors, which are involved in viral replication and cellular oncogenesis (3); however, their expression dynamics and the roles in each phase are poorly understood.

HTLV-1 is a human retrovirus that chronically infects $CD4^+$ T cells. Some infected individuals develop a malignant disease of $CD4^+CD25^+$ T cells, adult T-cell leukemia-lymphoma (ATL), and/or inflammatory diseases such as HTLV-1-associated myelopathy/tropical spastic paraparesis (HAM/TSP) and HTLV-1 uveitis (4, 5). Unlike that of human immunodeficiency virus (HIV), the replication of HTLV-1 is at a very low level *in vivo*, and viral RNA is rarely detected in the plasma of infected individuals (6). Instead, this virus persists in the host by using two different strategies: cell-to-cell transmission of viral particles (*de novo* infection) and clonal proliferation of infected cells (mitotic expansion) (7, 8).

HTLV-1 encodes two oncogenic factors, Tax and HTLV-1 bZIP factor (HBZ), in the sense and antisense strands of provirus respectively, and these two factors counteract each other in many signaling pathways (9). Tax is not only a potent oncoprotein (10) but also an efficient transactivator of viral replication, which means that Tax is required for *de novo* infection. However, Tax expression is generally suppressed in infected cells *in vivo* (11). In contrast, HBZ is constitutively expressed in infected cells and plays important roles in viral latency (8) and proliferation of infected cells (12). These facts suggest that HTLV-1 fine-tunes the expression and function of these counteracting viral factors to establish persistent infection in infected individuals. Since Tax is highly immunogenic (13-15), HTLV-1 appears to minimize Tax expression to escape from host immunity. Therefore, the significance of Tax in leukemogenesis has been an important unsolved issue.

In this study, we show that only a small fraction (0.05-3%) of cells transiently express Tax. However, knockdown of Tax induced apoptosis in the majority of cells, suggesting that Tax expressed in a minor subset is required for survival of the whole population. At a single-cell level, Tax-expressing cells highly expressed several anti-apoptotic and NF-κB-related genes, and Tax-negative cells were divided into two subpopulations which expressed medium or low level of the anti-apoptotic genes. Computational simulations support our hypothesis that transient Tax expression confers an anti-apoptotic property on the expressing cell, and that this effect lasts after Tax expression is diminished. This study is the first to show the dynamics of Tax expression at a single-cell level, and demonstrates novel roles of Tax in establishing persistent

infection by HTLV-1.

Results

A small fraction of MT-1 cells express Tax, and this expression is critical for the survival of the whole cell population

HTLV-1 *tax* is generally silenced or transcribed at quite low levels in ATL cells *in vivo* (16). An ATL cell line, MT-1, has an equivalent expression profile of viral genes to primary HTLV-1-infected cells (17). We carried out single-cell quantitative RT-PCR to elucidate the expression levels of *tax* and *HBZ* in individual MT-1 cells. The initial experiment showed that only one cell in 71 cells expressed a high level of *tax*, while no detectable expression was observed in remaining 70 cells (Fig. 1*A*). In contrast, *HBZ* was expressed in Tax-negative cells while it was not detected in the Tax-expressing cell. These results suggest that the expression of *tax* and *HBZ* is strictly and reciprocally regulated in MT-1 cells.

Our next question was, what is the role of Tax expression in a small number of MT-1 cells? To address this issue, we knocked down Tax with shRNA. Two different lentivirus vectors expressing shRNA targeting Tax (shTax1 and shTax4) were transduced into MT-1 cells, and either construct efficiently inhibited Tax expression to ~15% of the level in control cells (Fig. 1*B*, left). Intriguingly, progressive cell death was observed in MT-1 cells after Tax was knocked down (Fig. 1*B*, right), and as a mechanism, we found that Tax knockdown (KD) induced apoptosis of MT-1 cell, but not Jurkat cells (Fig. 1*C* and Fig. S1). To compare population dynamics between

Tax-KD cells and Tax-intact cells, a GFP competition assay was carried out (Fig. 1*D*, left schema). To our surprise, ~90% of shTax-transduced cells were eliminated from culture by day 21 (Fig. 1*D*, right), which was faster than we expected. These results indicated that Tax is indispensable for the survival of MT-1 cells even though only a small fraction of MT-1 cells express it. Since shTax-transduced cells were not rescued by the presence of non-transduced cells, it also appears that a cell's survival may depend at least partly on Tax expression within that cell, rather than depending on Tax expression by neighboring cells.

Tax is transiently expressed in MT-1 cells

We wished to further distinguish between these two possible hypotheses: one, that a small population of cells constantly expresses Tax and supports the survival of the whole population, or two, that all cells transiently express Tax by turns. To monitor Tax expression, we established a reporter subline of MT-1, MT1GFP, which expresses destabilized enhanced green fluorescent protein (d2EGFP; the half-life of the modified EGFP is 2 hours) under the control of 18 copies of the Tax-responsive element (TRE) (Fig. 1*E*, upper schema) (18). It was confirmed that d2EGFP expression in MT1GFP cells was closely correlated with Tax expression by intracellular staining with anti-Tax antibody, and a very small population (~0.5%) of MT1GFP cells expressed d2EGFP (Fig. 1*E*, bottom plot). Several stable clones of MT1GFP were established, and all clones possessed a small Tax-expressing subset. Using one of the MT1GFP clones, we evaluated the dynamics of Tax expression in individual cells by a time-lapse imaging.

The result revealed that d2EGFP expression was transient in most d2EGFP-positive cells (Fig. 1*F* and Movie S1), indicating that Tax is expressed temporarily in MT-1 cells. We could also see cells with fluctuating and continuous patterns of d2EGFP expression (see legend to Fig. 1*G* for definitions of these terms), although the percentages of cells with those patterns were lower than cells with transient expression (6% for fluctuating and 18% for continuous cells; Fig. 1*G*). Analysis of 87 cells with transient d2EGFP expression showed that the median and mean duration of Tax expression was approximately 22 and 18 hours, respectively (95% CI, 19.3, 24.7) (Fig. 1*H*).

It has been reported that two other ATL cell lines, KK-1 and SO-4, have expression levels of Tax and HBZ similar to those of fresh ATL cells (17). A small fraction of KK-1 and SO-4 cells each expressed Tax, and a subline of KK-1 reporting Tax expression, KK1GFP, was successfully established (Fig. S2*A* and *B*). Since DNA hyper-methylation in 5' long terminal repeat (LTR) is known to suppress *tax* transcription, we evaluated its methylation level and *tax* expression in fresh ATL cells and three ATL cell lines (MT-1, KK-1, and SO-4). In approximately half of ATL cases and all cell lines, DNA methylation level of 5'LTR was low and *tax* mRNA was detectable, suggesting that *tax* expression is inducible in these cells (Fig. S2*C*).

Computer simulation represents the dynamics of Tax expression

It was an incomprehensible observation that knockdown of Tax in a minor fraction (~0.5%) of MT-1 cells induced progressive apoptosis. To estimate the duration that all MT-1 cells experience Tax expression, we employed a computational simulation based

on a mathematical model. A previous study on HIV elegantly established a mathematical model of 5'LTR activation by Tat in latently infected cells (19). We adapted this model to simulate the regulation of the HTLV-1 5'LTR by Tax (See Fig. S3*A* for a schematic representation and Note S1 for details.) We tested several parameters used in the HIV study and adjusted them to fit our experimental data of transient Tax expression (Fig. S3*B-D*, Table S1, and Note S2). The simulation could then successfully reproduce the dynamics of Tax expression in MT-1 cells, and the calculated mean duration of Tax expression was 13.6 hours, consistent with experimental observations (Fig. S3*C*), suggesting that the parameters we used were suitable for the model for activation of the HTLV-1 LTR. Using this model, we then estimated the distribution of the interval between two successive Tax expression episodes (Fig. S3*E*), and interestingly, our simulation reproduced the single-cell level expression pattern of Tax (Fig. S3*F*). Furthermore, we calculate that it takes approximately 150 days for 90% of MT-1 cells to experience Tax expression (Fig. S3*G*). This result means that the loss of Tax-KD cells occurs much faster than the Tax expression (Fig. 1*D*). The details of our single-cell level computational simulations are described in the SI Notes.

This simulation suggests a possible mechanism that short-term expression of Tax by a cell confers some effects on not only to the cell during its period of Tax expression but also to the cell and its progeny even after Tax expression is lost.

Distinct transcriptional profiles in Tax-expressing cells

To clarify the effects of Tax on both Tax-expressing cells and non-expressing cells, we analyzed the transcriptional profiles of each population. Tax-expressing MT1GFP or KK1GFP cells (Fig. 1*E* or Fig. S2*B*, respectively) were purified from bulk cells by a cell sorter using d2EGFP as a marker for Tax, and RNAs from each fraction were subjected to RNA-Seq. Many T-cell activation-associated genes were differentially regulated in Tax-expressing cells compared to non-expressing cells (Fig. 2*A* and *B*). Upstream regulator analysis was carried out using the Ingenuity Pathway Analysis program and revealed that NF-κB-related pathways were significantly affected by Tax in both cell lines (Fig. 2*B*).

To analyze the correlation between Tax expression and that of other genes at a single-cell level, we chose ~90 representative genes in addition to *HBZ*, and compared their mRNA levels in sorted Tax-positive vs. Tax-negative MT-1 cells by single-cell quantitative RT-PCR. As expected, clustering analysis could clearly separate Tax-positive cells from negative cells (Fig. 2*C*). Expression of NF-κB target genes and apoptosis-related genes was positively correlated with that of *tax* in each MT-1 cell (Fig. 2*D* and Fig. S4). Interestingly, violin plots of several anti-apoptotic genes, such as *CFLAR*, *GADD45B*, *TRAF1*, and *TNFAIP3*, showed 2 peaks in Tax-negative cells; one with almost no expression and another with lower expression than Tax-positive cells (Fig. 2*D*). In addition, a 3D principal component analysis (PCA) plot generated by the expression levels of these genes revealed that there were 2 clusters in Tax-negative cells (Fig. 2*E*). These findings suggest that Tax influences the expression of anti-apoptotic genes in Tax-negative cells, there are 2 subpopulations -- those with medium and low

levels of anti-apoptotic factors and thus different degrees of sensitivity to apoptosis.

Tax expression is induced by cytotoxic stresses and contributes an anti-apoptotic property to cells

Since Tax expression was associated with the upregulation of anti-apoptotic genes, we hypothesized that Tax might be induced in response to cytotoxic stress. First, we evaluated the effect of the phases of cell growth on Tax expression. As shown in Fig. 3*A*, the percentage of d2EGFP-positive (Tax-expressing) cells increased stepwise from 1.2% at day 1 to 11.5% at day 7, if cells were cultured without passage. In contrast, the ratio of d2EGFP-expressing cells remained less than 3% when cells were properly passed, suggesting that stresses associated with the cell overgrowth triggered Tax expression. Importantly, under stressed conditions, the viability of Tax-expressing (d2EGFP+) cells was strikingly higher than that of Tax-negative (d2EGFP-) cells (Fig. 3*B*). These observations suggest that stress-induced Tax expression confers anti-apoptotic capacities on MT-1 cells.

Similarly, an inducer of oxidative stress, H_2O_2, induced Tax (Fig. 3*C*, left), and the viability of Tax-expressing cells was significantly higher than that of non-expressing cells (Fig. 3*C*, right). It has been reported that one of the platinum-containing anti-cancer drugs, cisplatin, induces production of reactive oxygen species (ROS) in cancer cells and contributes to cytotoxicity (20). Cisplatin induced Tax in MT1GFP cells, while an inhibitor of ROS, N-acetyl-L-cysteine (NAC), completely canceled its effect (Fig. 3*D*). These results suggest that Tax functions as a safeguard against

apoptosis induced by various cytotoxic stresses.

Similar mechanisms for latency of HTLV-1 and HIV

It has been known that cytotoxic stresses, such as DNA damage and oxidative stress, activate the HIV LTR (21, 22). Several HIV reactivating reagents (disulfiram, panobinostat, SAHA, and JQ-1) could also induce Tax expression in MT1GFP cells (Fig. 3*E*). JQ-1 is an inhibitor of bromodomain-containing 4 (BRD4) and robustly reactivates HIV transcription when it is used with phorbol 12-myristate 13-acetate (PMA) and ionomycin (23). A combination of JQ-1 and PMA/ionomycin exhibited potent activity on Tax induction in MT1GFP cells (Fig. 3*F*), suggesting overlapping mechanisms for latency of HTLV-1 and HIV.

It has been reported that minimum feedback circuit between HIV Tat and the HIV 5'LTR is sufficient to establish HIV latency (19). When a lentivirus vector expressing Tat through the HIV 5'LTR was transduced into Jurkat cells, gene expression from the 5'LTR was highly transactivated in most cells, while a small population of infected cells became latent (24). Since our results suggested that Tax has roles similar to Tat in latency, it was tested if the HTLV-1 LTR-Tax circuit can generate latency in T cells. We found that in Jurkat cells expressing Tax and d2EGFP through the HTLV-1 5'LTR (Jurkat/LTRd2EGFP-Tax cells), a very small population expressed d2EGFP -- just as was the case for MT1GFP cells (Fig. S5*A*, right). In contrast, a similar construct containing HIV Tat and HIV LTRs induced continuous expression of Tat and d2EGFP in more than 95% of cells, while a remaining subset was dormant (Fig. S5*B*); this

observation was compatible with the previous report (24). Treatment by JQ-1, PMA and ionomycin together could activate expression of d2EGFP in all Jurkat/LTRd2EGFP-Tax clones (Fig. S5*C*), and time-lapse imaging revealed transient expression of d2EGFP in some untreated Jurkat/LTRd2EGFP-Tax cells (Fig. S5*D*). These findings suggest that the Tax-5'LTR circuit is a basic unit for modulating Tax expression flexibly in response to various stimulations.

Tax suppresses the cell cycle transition from S to G2/M

A possible alternate explanation for the big impact of Tax KD on MT-1 cell survival was that Tax-expressing cells might proliferate faster than non-expressing cells. To test this hypothesis, the effect of Tax on the cell cycle was analyzed using a method combining DAPI staining and EdU incorporation (Fig. 4*A*) -- a method by which we can evaluate temporal cell cycle progression without synchronization (25). We found that the proportion of cells in S and G2/M phases two hours after the start of an EdU pulse was comparable between Tax-positive and Tax-negative cells. But at 6 hours, the proportion of G2/M cells within the Tax-positive population was only ~80% of that within the Tax-negative population (G2/M cells comprised 39.4% and 48.7% of Tax-positive and negative cells, respectively), suggesting that if anything, Tax retarded, rather than sped up, the G2/M transition of MT-1 cells (Fig. 4*B* and *C*). This result implies that Tax does not promote faster proliferation of expressing cells; rather, the anti-apoptotic property of Tax is responsible for the survival of MT-1 cells.

The carryover of anti-apoptotic factors into the Tax-negative interval is important for maintaining the MT-1 cell population

Based on the results of the Tax KD and single-cell analysis experiments, we hypothesized that transient Tax expression triggers anti-apoptotic genes in a small number of MT-1 cells, that this effect is carried over in a significant number of these cells when they enter the Tax-negative phase, and that such carryover is critical for the survival and expansion of the cell population. To check this hypothesis, we constructed a population level agent-based model (ABM) that can represent our experimental observations. More precisely, the ABM simulates the population dynamics of MT-1 cells transduced with shNC or shTax4 as an inhomogeneous (birth-death) Poisson process with a stage transition (See Fig. 5*A* for a schematic representation and Note S3 for details.). We carried out ABM simulations and confirmed that the time-course of experimental data for the number and fraction of shNC and shTax4 cells were well reproduced in Fig. 5*B* and *C*. Based on our ABM simulations, we reconstructed the time-course of the frequency of Tax-positive cells under normal conditions. As shown in Fig. 5*D*, our simulation predicted that a small number of MT-1 cells would express Tax (i.e., around 3% at the steady state). In addition, we calculated the predicted subpopulation dynamics of shNC and shTax4 cells (Fig. 5*E* and *F*, respectively). The numbers of cells that are Tax positive and anti-apoptotic gene positive (i.e., T_{on}), Tax negative and anti-apoptotic gene high (i.e., $T_{\mathrm{off}}A_{\mathrm{high}}$), and Tax negative and anti-apoptotic gene low (i.e., $T_{\mathrm{off}}A_{\mathrm{low}}$) are shown. Under normal conditions (Fig. 5*E)*, the fraction of cells that are T_{on} remains small, but the number of cells in each

subpopulation increases and the total cell population expands, since MT-1 cells are able to express Tax and this expression protects them from apoptosis. In fact, our simulation predicts that the major subpopulation will be $T_{\mathrm{off}}A_{\mathrm{high}}$, which is consistent with our detection of the expression of apoptosis related genes in Fig. 2*D*. On the other hand, under Tax knockdown conditions (Fig. 5*F*), the fraction of cells that are T_{on} decreases until there are almost none left, because there is no further transition of cells from Tax negative to Tax positive. The collapse of the T_{on} subpopulation leads to a corresponding decrease in anti-apoptotic gene expression, renders more cells susceptible to apoptosis, and decreases the degree to which the total cell population expands.

To examine whether having experienced Tax expression prolongs cell survival time, we also calculated the survival probabilities of the subpopulations of shNC cells in our ABM simulations. (See Note S4 for details.) For this calculation, we defined the subpopulations of shNC cells based on whether Tax had been expressed at least once or had never been expressed (solid and dashed lines in Fig. 5*G*, respectively). Interestingly, in this simulation, Tax expression significantly increases the average lifespan of the cell.

Taken together, these simulations support a novel model for ATL persistence: transient expression of Tax in cells is responsible for cell population-level maintenance and expansion (Fig. 6). In a steady state (Fig. 6, left), T_{on} cells turn successively into $T_{\mathrm{off}}A_{\mathrm{high}}$ and then $T_{\mathrm{off}}A_{\mathrm{low}}$ cells. Although $T_{\mathrm{off}}A_{\mathrm{low}}$ cells are sensitive to apoptotic signals and gradually die in culture, Tax-negative cells proliferate more rapidly than Tax-positive cells. In response to cytotoxic stresses, Tax can be induced in some

Tax-negative cells, and thus T_{on} cells are replenished, a process which is critical for the persistence of the population. Tax knockdown (Fig. 6, right) causes a decrease in T_{on} cells, resulting in a shortage of apoptosis-resistant cells and the induction of massive apoptosis. This model enables us to explain how such a small number of Tax-expressing cells has a big impact on the dynamics of the whole population.

Discussion

Persistent viruses have evolved shrewd strategies to propagate *in vivo* while evading host immune surveillance. Here, we demonstrate that HTLV-1 utilizes a unique way to enhance survival and proliferation of infected cells: the transient expression of Tax confers an anti-apoptotic property to cells and maintains the whole population. Since it has been reported that continuous expression of Tax induces DNA damage and senescence in the expressing cells (26, 27), it is suggested that short-term expression of Tax is beneficial to cells. Interestingly, a similar phenomenon of cell survival being promoted by transient gene expression was reported in mouse embryonic stem cells (mESCs). *Zscan4*, which functions in the maintenance of telomeres, is transiently expressed in only ~5% of mESCs at any given time. Knockdown of *Zscan4* induces massive cell death before all cells experience its expression (28). Moreover, expression of *Zscan4* is linked to activation of an endogenous retrovirus, MERVL (29). Those studies and our observations suggest an association between the transient activation of retroviral LTRs and the maintenance of cell populations.

Tax expression is essential for *de novo* infection by HTLV-1, since viral transcription

depends on Tax (30, 31). However, Tax expression strongly induces expression of viral proteins, including Tax, Env and Gag, resulting in attacks by cytotoxic T lymphocytes (CTLs). Therefore, intermittent Tax expression is a clever strategy of HTLV-1 to evade the host immune response most of the time, yet maintain the ability to cause *de novo* infection under certain conditions. A recent study has reported that *tax* is induced by hypoxia (32); it is compatible with the previous observation that high *tax* expression was detected in the bone marrow, which is physiologically hypoxic (33). As another example, HTLV-1 can be transmitted through breast-feeding -- a process in which HTLV-1 infected cells have to pass through the alimentary tract with acidic conditions and bile acids. Stress-induced Tax expression would be beneficial for *de novo* infection in these conditions. It is known that low pH and hypoxia in the physiological environment suppress adaptive immunity (34, 35), suggesting that infected cells may be able to "get away with" expressing Tax for a limited time in such immunological niches. To clarify the *in vivo* dynamics of Tax expression in immunocompetent hosts, further studies using animal models will be required.

Tax expression is suppressed by genetic/epigenetic mechanisms in 50% of ATL cases (16, 36). The other half of ATL cases have the potential to express Tax, and indeed, it is known that nearly 50% of ATL cases express viral antigens after *ex vivo* culture (37). In this study, we show that survival of MT-1 cells, which express a level of *tax* similar to that of primary ATL cells (17), depends on Tax, suggesting that clinical ATL cases are divided into at least two classes: Tax-dependent and -independent ATL. Previous finding that the arsenic/interferon treatment induced apoptosis of several Tax-expressing cell

lines including MT-1 through degradation of Tax protein implies that those established cell lines still possess the characteristics of Tax-dependent ATL (38). It is known that the activity of Tax-specific CTLs is an important factor in inhibiting the onset of ATL (15). A recent study provided more evidence that Tax plays an important role in ATL: a clinical trial of a dendritic cell vaccine targeting Tax was efficacious (39). Tax expression was faint in fresh ATL cells of the enrolled patients, but it was inducible by *ex vivo* culture. These reports support our hypothesis that transiently expressed Tax plays critical roles in the development and maintenance of ATL even if its expression is limited to a small fraction of leukemic cells. Recent comprehensive genomic studies of clinical ATL cases showed that genetic/epigenetic alterations accumulate in Tax-associated genes (40, 41), suggesting that the effects of these changes can substitute for Tax functions in leukemogenesis.

It is noteworthy that expression of *tax* was contrary to that of *HBZ* even at the single cell level ($r=-0.8$, $p<0.0001$, Fig. 2*D*). A similar observation in HTLV-1-infected T-cell clones was published while our study was under consideration for publication (42). That study showed the heterogeneity in the expression levels of *tax* and *HBZ* mRNA in fixed cells, whereas we here demonstrate the dynamics of Tax expression using live-cell imaging. Tax and HBZ have opposing functions in many signaling pathways (9); however, the significance of their contrary activities has not been clarified. In this study, we found that Tax induces anti-apoptotic genes, such as *CFLAR*, *TNFRSF9*, *TRAF1*, and *TNFAIP3* (43-45). In contrast, pathways associated with cell proliferation, such as CD3, CD28, and melanogenesis associated transcription factor (MITF) (46, 47), are

significantly activated in Tax-negative cells that express HBZ (Fig. 2*B*). This finding is compatible with the previous studies that showed that HBZ promotes T-cell growth by enhancing signals through CD3, and that knockdown of HBZ suppresses the proliferation of ATL cells including MT-1 cells (12, 48, 49). The alternation of expression of Tax and HBZ seems to execute cooperative programs for survival and proliferation rather than causing them to interfere with each other. It has been reported that hyper-activation of the NF-κB pathway by Tax triggers senescence in HeLa cells, while HBZ could cancel this suppressive effect (27), suggesting that the collaboration of Tax and HBZ is important for the expansion of HTLV-1-infected cells. Transgenic mouse models have demonstrated that both Tax and HBZ are oncogenic (50); however, their roles in HTLV-1-infected individuals are thought to be more complicated, due to immune surveillance and the fact that human cells are more resistant to malignant transformation than rodent cells (51, 52). HBZ plays important roles in determining the immunophenotype of HTLV-1-infected cells, and HBZ itself has low immunogenicity (53, 54). It has been reported that the expression level of *HBZ* is higher in ATL cells than in non-leukemic infected cells, implying that infected clones with higher *HBZ* expression are selected during leukemogenesis (17). These findings suggest that constant expression of HBZ drives proliferation of cells with this specific immunophenotype, while transient Tax expression engages to inhibit of cell death caused by various stresses. The different functions and expression patterns of Tax vs. HBZ are thought to be important in the malignant transformation of human T cells.

Recent studies show that many pathogens can persist in their reservoirs during both

acute and chronic infections, and re-emerge in the host under stressful conditions (55). In this study, we demonstrate that the HTLV-1 5'LTR is activated by various cytotoxic stresses and several reagents that reactivate latent HIV (Fig. 3). These results suggest that HTLV-1-infected cells utilize mechanisms analogous to those of latent HIV-infected cells to act as reservoirs. One example is the antagonism of Tax and Tat to BRD4. BRD4 is a negative regulator of positive transcription elongation factor b (P-TEFb). It has been reported that both Tax and Tat compete with BRD4 for binding to P-TEFb (56, 57), and indeed, we found that a BRD4 inhibitor, JQ-1, efficiently reactivated the HTLV-1 5'LTR (Fig. 3*E* and *F*, and Fig. S5*C*). These findings suggested that BRD4 can interrupt the positive feedback loop generated by the 5'LTR and Tax/Tat, and allow the establishment of latent infection by HTLV-1/HIV. Interestingly, many other viruses, including KSHV and human papilloma virus, also use BRD4 and/or P-TEFb to regulate viral life cycles and pathogenesis (58-60). Identification of the cellular factors acted on by diverse viruses will contribute to exploring common mechanisms of viral latency and reactivation. In the case of HTLV-1, several repressors of viral replication, such as p30, Rex, and HBZ are encoded in the provirus (9). These viral factors may be able to regulate the duration and/or timing of Tax expression in infected cells more elaborately. Additional investigation will be required to understand the precise roles of each protein in HTLV-1 latency.

Since the discovery of HTLV-1, a number of studies on its pathogenesis have been conducted; however, the prevention and treatment of HTLV-1-induced diseases are still unsatisfactory (61). In this study, we demonstrated that sporadic and transient Tax

expression in a small subset of cells has significant influence on gene expression of their progeny cells, and consequently maintains the whole population of ATL cells. This is a novel mechanism of the retrovirus for persistence and latency *in vivo*, and elucidation of this mechanism can contribute to a better understanding of viral pathogenesis and the development of new strategies for treatment and prophylaxis of viral-associated diseases.

Materials and Methods

Cells

An IL-2 independent ATL cell line, MT-1 (62), two IL-2 dependent ATL cell lines, KK1 and SO4 (63), and an HTLV-1 negative T-cell line, Jurkat, were used in this study. MT-1 and Jurkat cells were maintained in RPMI supplemented with 10% (vol/vol) FBS. The MT1GFP cell line was maintained as MT-1 cells were, with the addition of 500 ug/ml G418 (Nacalai). KK1 and SO4 cells lines were maintained in RPMI supplemented with 10% FBS and IL-2 (100 U/ml) (PeproTech).

Clinical samples

Fresh ATL cells were obtained from 20 aggressive type ATL cases, and used for extraction of genomic DNA and total RNA. Use of the clinical samples in this research was approved by the ethics committee of Kyoto University (approval number: G204). Written consent was obtained from the patients. Using genomic DNAs from primary ATL cells and ATL cell lines, DNA methylation level of 5'LTR was analyzed by Combined Bisulfite Restriction Analysis (COBRA) method as previously described (36).

Expression level of *tax* in fresh ATL cells was analyzed by a conventional quantitative RT-PCR (16).

GFP Competition assay

The GFP competition assay (64) was carried out to observe the long term effect of Tax knockdown on MT-1 or Jurkat cells transduced with pLKO-GFP lentivirus expressing shNC, shTax1 or shTax4. Cells were infected with concentrated lentivirus at an MOI of 0.5 to adjust the ratio of transduced cells to around 50%. The effect of shRNAs on target cells was evaluated by measuring the percentage of GFP+ cells using a FACSverse flow cytometer (BD Biosciences).

Single cell quantitative PCR (qPCR)

The C1 Single-Cell Auto Prep Array for PreAmp (Fluidigm) was used for harvest of RNA, cDNA synthesis, and pre-amplification of cDNA (18 cycles of PCR for the target genes) from single cells according to manufacturer's instructions. After loading cells onto an integrated fluidic circuit (IFC), we checked all 96 chambers by microscope to verify capture of a single cell. Thereafter, pre-amplified cDNA was harvested and subjected to qPCR. The Biomark HD system (Fluidigm) combined with EvaGreen chemistry (Bio-Rad) was used for the qPCR assay. To increase specificity, we used nested primers (one pair for the pre-amplification step and another pair for the subsequent qPCR of 30 cycles). The sequences of the primers used in this study are indicated in Table S3. Raw data was processed by Fluidigm Real-Time PCR analysis software and the melting curve was used to determine the Pass/Failure call of qPCR. Data was analyzed with the R program using the Singular Analysis Toolset package

(Fluidigm). Any chamber which contained more than one cell was excluded from analysis; outlier cells that had low global expression were also excluded. The level of detection value was set to 24 cycles according to manufacturer's recommendation.

Time lapse imaging

For live cell imaging, $8x10^4$ MT1GFP cells were seeded in a 5 mm glass bottom dish (Matsunami) pre-coated with Poly-D-lysine (Sigma), and incubated at $37^{O}C$ in 5% CO_2. Images in the DIC and GFP channels were captured with an LCV110 microscope (Olympus) every 20 minutes for 96 hours. Semi-automated cell tracking was done by Fiji software with the Trackmate plugin (65). Cells, which had already expressed d2EGFP at beginning of observation, were excluded from analysis because the starting point for expression was unknown. To analyze single-cell dynamics of d2EGFP expression, normalized fluorescence intensities are plotted against time. The starting time (t=0) is the time at which the cell started expressing d2EGFP above background level.

RNA-seq

MT1GFP or KK1GFP cells were sorted into d2EGFP+ and d2EGFP- populations with a FACSAria III (BD Biosciences), and RNA was then extracted using the RNeasy mini kit (Qiagen). Single-end RNA sequencing was performed (BGI). A quality check was done with FastQC, and then Tuxedo pipeline was used for RNA quantification (66). Upstream regulator analysis was carried out by Ingenuity Pathway Analysis (Qiagen).

Cell cycle analysis

To measure cell the cycle dynamics of MT1GFP cells without cell synchronization, a

method combining Edu incorporation and DAPI staining was carried out as previously described (25) using the Click-iT Plus EdU Alexa Fluor 647 kit (Thermo Fisher). Initially $1x10^7$ cells were cultured in complete RPMI medium containing 10 μM Edu for 2 hours and washed thereafter. At that point (time=2 hours), half of the cells were fixed and stained for Edu and DAPI according to the manufacturer's protocol. The other half of the cells were re-cultured again for 4 hours without adding Edu. At the end of the assay (time=6 hours) those cells were washed, fixed and stained. Cell cycle status in the d2EGFP+ or d2EGFP- population was analyzed using a FACSVerse based on the levels of Edu and DAPI.

Statistical analysis

Statistical analysis was done using Microsoft Excel, GraphPad Prism, R or Python. Data obtained by flow cytometry was analyzed with FlowJo. Two-sided t-test was used to compare between two different groups.

Computational simulations

A simple deterministic two-state model of Tax positive feedback was developed (see the chemical reaction scheme in Fig. S3*A*) (19), and simulated by the Gillespie algorithm (67) to investigate the stochastic dynamics of Tax expression, especially to estimate the length of Tax expression (t_{period}) and the interval between Tax expression episodes (t_{interval}). In addition, an agent-based model (ABM) of cell population dynamics (birth-death Poisson process) with stage transitions as described in Fig. 5*A* was constructed based on the individual-based Gillespie algorithm to confirm the experiment-based hypothesis that the transient expression of Tax is a critical event for

the persistence of the MT-1 cell population. See SI Notes for further details.

Acknowledgements

We thank Dr. Mitsuaki Yoshida (The Cancer Institute, Japanese Foundation For Cancer Research) for helpful discussion, Dr. Chou-Zen Giam (Uniformed Services University of the Health Sciences) for providing PB-18X21-RFP and PB-Tase, Drs. Hiroo Hasegawa and Yasuaki Yamada (Nagasaki University) for KK1 and SO4, and Dr. Linda Kingsbury for proofreading. This study was supported by the Project for Cancer Research And Therapeutic Evolution (P-CREATE) (17cm0106306h0002 to J.Y. and M.M.), the Research Program on Emerging and Re-emerging Infectious Diseases (17fk0108227h0002 to J.Y. and M.M.), and J-PRIDE (17fm0208006h0001, 17fm0208019h0101, and 17fm0208014h0001 to S.I.) from Japan Agency for Medical Research and development (AMED), JSPS KAKENHI (JP16H05336 to M.M. and JP17K07166 to J.Y.), the JST PRESTO and CREST program (to S.I.), a Grant-in-Aid for Scientific Research on Innovative Areas from MEXT (16H06429, 16K21723, and 17H05819 to S.I.), a grant from Princess Takamatsu Cancer Research Fund (to J.Y.), and a grant from The Yasuda Medical Foundation (to J.Y.). This study was also supported in part by the JSPS Core-to-Core Program A, Advanced Research Networks.

References

1. Virgin HW, Wherry EJ, & Ahmed R (2009) Redefining chronic viral infection. *Cell* 138:30-50.
2. Traylen CM*, et al.* (2011) Virus reactivation: a panoramic view in human infections. *Future Virol* 6:451-463.
3. Mesri EA, Feitelson MA, & Munger K (2014) Human viral oncogenesis: a cancer hallmarks analysis. *Cell Host Microbe* 15:266-282.
4. Tagaya Y & Gallo RC (2017) The Exceptional Oncogenicity of HTLV-1. *Front Microbiol* 8:1425.
5. Matsuoka M & Jeang KT (2007) Human T-cell leukaemia virus type 1 (HTLV-1) infectivity and cellular transformation. *Nat Rev Cancer* 7:270-280.
6. Demontis MA, Sadiq MT, Golz S, & Taylor GP (2015) HTLV-1 viral RNA is detected rarely in plasma of HTLV-1 infected subjects. *J Med Virol* 87:2130-2134.
7. Melamed A*, et al.* (2013) Genome-wide determinants of proviral targeting, clonal abundance and expression in natural HTLV-1 infection. *PLoS Pathog* 9:e1003271.
8. Philip S, Zahoor MA, Zhi H, Ho YK, & Giam CZ (2014) Regulation of human T-lymphotropic virus type I latency and reactivation by HBZ and Rex. *PLoS Pathog* 10:e1004040.
9. Matsuoka M & Yasunaga J (2013) Human T-cell leukemia virus type 1: replication, proliferation and propagation by Tax and HTLV-1 bZIP factor. *Curr Opin Virol* 3:684-691.
10. Boxus M*, et al.* (2008) The HTLV-1 Tax interactome. *Retrovirology* 5:76.
11. Hanon E*, et al.* (2000) Abundant tax protein expression in CD4+ T cells infected with human T-cell lymphotropic virus type I (HTLV-I) is prevented by cytotoxic T lymphocytes. *Blood* 95:1386-1392.
12. Satou Y, Yasunaga J, Yoshida M, & Matsuoka M (2006) HTLV-I basic leucine zipper factor gene mRNA supports proliferation of adult T cell leukemia cells. *Proc Natl Acad Sci U S A* 103:720-725.
13. Jacobson S, Shida H, McFarlin DE, Fauci AS, & Koenig S (1990) Circulating

CD8+ cytotoxic T lymphocytes specific for HTLV-I pX in patients with HTLV-I associated neurological disease. *Nature* 348:245-248.

14. Kannagi M*, et al.* (1991) Predominant recognition of human T cell leukemia virus type I (HTLV-I) pX gene products by human CD8+ cytotoxic T cells directed against HTLV-I-infected cells. *Int Immunol* 3:761-767.
15. Kannagi M, Hasegawa A, Takamori A, Kinpara S, & Utsunomiya A (2012) The roles of acquired and innate immunity in human T-cell leukemia virus type 1-mediated diseases. *Front Microbiol* 3:323.
16. Takeda S*, et al.* (2004) Genetic and epigenetic inactivation of tax gene in adult T-cell leukemia cells. *Int J Cancer* 109:559-567.
17. Usui T*, et al.* (2008) Characteristic expression of HTLV-1 basic zipper factor (HBZ) transcripts in HTLV-1 provirus-positive cells. *Retrovirology* 5:34.
18. Zhang L, Liu M, Merling R, & Giam CZ (2006) Versatile reporter systems show that transactivation by human T-cell leukemia virus type 1 Tax occurs independently of chromatin remodeling factor BRG1. *J Virol* 80:7459-7468.
19. Razooky BS, Pai A, Aull K, Rouzine IM, & Weinberger LS (2015) A hardwired HIV latency program. *Cell* 160:990-1001.
20. Marullo R*, et al.* (2013) Cisplatin induces a mitochondrial-ROS response that contributes to cytotoxicity depending on mitochondrial redox status and bioenergetic functions. *PLoS One* 8:e81162.
21. Valerie K*, et al.* (1988) Activation of human immunodeficiency virus type 1 by DNA damage in human cells. *Nature* 333:78-81.
22. Legrand-Poels S, Vaira D, Pincemail J, van de Vorst A, & Piette J (1990) Activation of human immunodeficiency virus type 1 by oxidative stress. *AIDS Res Hum Retroviruses* 6:1389-1397.
23. Zhu J*, et al.* (2012) Reactivation of latent HIV-1 by inhibition of BRD4. *Cell Rep* 2:807-816.
24. Weinberger LS, Burnett JC, Toettcher JE, Arkin AP, & Schaffer DV (2005) Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. *Cell* 122:169-182.
25. Fleisig H & Wong J (2012) Measuring cell cycle progression kinetics with metabolic labeling and flow cytometry. *J Vis Exp*:e4045.

26. Kinjo T, Ham-Terhune J, Peloponese JM, Jr., & Jeang KT (2010) Induction of reactive oxygen species by human T-cell leukemia virus type 1 tax correlates with DNA damage and expression of cellular senescence marker. *J Virol* 84:5431-5437.
27. Zhi H*, et al.* (2011) NF-kappaB hyper-activation by HTLV-1 tax induces cellular senescence, but can be alleviated by the viral anti-sense protein HBZ. *PLoS Pathog* 7:e1002025.
28. Zalzman M*, et al.* (2010) Zscan4 regulates telomere elongation and genomic stability in ES cells. *Nature* 464:858-863.
29. Eckersley-Maslin MA*, et al.* (2016) MERVL/Zscan4 Network Activation Results in Transient Genome-wide DNA Demethylation of mESCs. *Cell reports* 17:179-192.
30. Cann AJ, Rosenblatt JD, Wachsman W, Shah NP, & Chen IS (1985) Identification of the gene responsible for human T-cell leukaemia virus transcriptional regulation. *Nature* 318:571-574.
31. Felber BK, Paskalis H, Kleinman-Ewing C, Wong-Staal F, & Pavlakis GN (1985) The pX protein of HTLV-I is a transcriptional activator of its long terminal repeats. *Science* 229:675-679.
32. Kulkarni A*, et al.* (2017) Glucose Metabolism and Oxygen Availability Govern Reactivation of the Latent Human Retrovirus HTLV-1. *Cell Chem Biol*.
33. Levin MC*, et al.* (1997) Extensive latent retroviral infection in bone marrow of patients with HTLV-I-associated neurologic disease. *Blood* 89:346-348.
34. Kareva I & Hahnfeldt P (2013) The emerging "hallmarks" of metabolic reprogramming and immune evasion: distinct or linked? *Cancer Res* 73:2737-2742.
35. Taylor CT & Colgan SP (2017) Regulation of immunity and inflammation by hypoxia in immunological niches. *Nat Rev Immunol*.
36. Taniguchi Y*, et al.* (2005) Silencing of human T-cell leukemia virus type I gene transcription by epigenetic mechanisms. *Retrovirology* 2:64.
37. Kurihara K*, et al.* (2005) Potential immunogenicity of adult T cell leukemia cells in vivo. *Int J Cancer* 114:257-267.
38. Dassouki Z*, et al.* (2015) ATL response to arsenic/interferon therapy is triggered

by SUMO/PML/RNF4-dependent Tax degradation. *Blood* 125:474-482.

39. Suehiro Y*, et al.* (2015) Clinical outcomes of a novel therapeutic vaccine with Tax peptide-pulsed dendritic cells for adult T cell leukaemia/lymphoma in a pilot study. *Br J Haematol* 169:356-367.
40. Kataoka K*, et al.* (2015) Integrated molecular analysis of adult T cell leukemia/lymphoma. *Nat Genet* 47:1304-1315.
41. Fujikawa D*, et al.* (2016) Polycomb-dependent epigenetic landscape in adult T-cell leukemia. *Blood* 127:1790-1802.
42. Billman MR, Rueda D, & Bangham CRM (2017) Single-cell heterogeneity and cell-cycle-related viral gene bursts in the human leukaemia virus HTLV-1. *Wellcome Open Res* 2:87.
43. Okamoto K, Fujisawa J, Reth M, & Yonehara S (2006) Human T-cell leukemia virus type-I oncoprotein Tax inhibits Fas-mediated apoptosis by inducing cellular FLIP through activation of NF-kappaB. *Genes Cells* 11:177-191.
44. Saitoh Y*, et al.* (2016) A20 targets caspase-8 and FADD to protect HTLV-I-infected cells. *Leukemia* 30:716-727.
45. Sabbagh L*, et al.* (2013) Leukocyte-specific protein 1 links TNF receptor-associated factor 1 to survival signaling downstream of 4-1BB in T cells. *J Leukoc Biol* 93:713-721.
46. Acuto O & Michel F (2003) CD28-mediated co-stimulation: a quantitative support for TCR signalling. *Nat Rev Immunol* 3:939-951.
47. Flaherty KT, Hodi FS, & Fisher DE (2012) From genes to drugs: targeted strategies for melanoma. *Nat Rev Cancer* 12:349-361.
48. Arnold J, Zimmerman B, Li M, Lairmore MD, & Green PL (2008) Human T-cell leukemia virus type-1 antisense-encoded gene, Hbz, promotes T-lymphocyte proliferation. *Blood* 112:3788-3797.
49. Kinosada H*, et al.* (2017) HTLV-1 bZIP Factor Enhances T-Cell Proliferation by Impeding the Suppressive Signaling of Co-inhibitory Receptors. *PLoS Pathog* 13:e1006120.
50. Yasunaga J & Matsuoka M (2011) Molecular mechanisms of HTLV-1 infection and pathogenesis. *Int J Hematol* 94:435-442.
51. Hahn WC*, et al.* (1999) Creation of human tumour cells with defined genetic

elements. *Nature* 400:464-468.

52. Grassmann R, Aboud M, & Jeang KT (2005) Molecular mechanisms of cellular transformation by HTLV-1 Tax. *Oncogene* 24:5976-5985.
53. Bangham CRM & Matsuoka M (2017) Human T-cell leukaemia virus type 1: parasitism and pathogenesis. *Philos Trans R Soc Lond B Biol Sci* 372.
54. Ma G, Yasunaga J, & Matsuoka M (2016) Multifaceted functions and roles of HBZ in HTLV-1 pathogenesis. *Retrovirology* 13:16.
55. Anonymous (2017) Lessons from reservoirs. *Nat Med* 23:899.
56. Bisgrove DA, Mahmoudi T, Henklein P, & Verdin E (2007) Conserved P-TEFb-interacting domain of BRD4 inhibits HIV transcription. *Proc Natl Acad Sci U S A* 104:13690-13695.
57. Cho WK*, et al.* (2007) Modulation of the Brd4/P-TEFb interaction by the human T-lymphotropic virus type 1 tax protein. *Journal of virology* 81:11179-11186.
58. Chen HS*, et al.* (2017) BET-Inhibitors Disrupt Rad21-Dependent Conformational Control of KSHV Latency. *PLoS Pathog* 13:e1006100.
59. Jang MK, Shen K, & McBride AA (2014) Papillomavirus genomes associate with BRD4 to replicate at fragile sites in the host genome. *PLoS Pathog* 10:e1004117.
60. Zaborowska J, Isa NF, & Murphy S (2016) P-TEFb goes viral. *Inside Cell* 1:106-116.
61. Willems L*, et al.* (2017) Reducing the global burden of HTLV-1 infection: An agenda for research and action. *Antiviral Res* 137:41-48.
62. Miyoshi I*, et al.* (1980) A novel T-cell line derived from adult T-cell leukemia. *Gan* 71:155-156.
63. Yamada Y*, et al.* (1998) Interleukin-15 (IL-15) can replace the IL-2 signal in IL-2-dependent adult T-cell leukemia (ATL) cell lines: expression of IL-15 receptor alpha on ATL cells. *Blood* 91:4265-4272.
64. Eekels JJ*, et al.* (2012) A competitive cell growth assay for the detection of subtle effects of gene transduction on cell proliferation. *Gene Ther* 19:1058-1064.
65. Jaqaman K*, et al.* (2008) Robust single-particle tracking in live-cell time-lapse sequences. *Nat Methods* 5:695-702.

66. Trapnell C*, et al.* (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 7:562-578.
67. Gillespie DT (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics* 22:403-434.

Author contributions

M. Mahgoub, J.Y., K.S., and M. Matsuoka designed, performed, and/or analyzed all experiments. S.I., S.N., and Y.K. performed computational simulations. M. Mahgoub, J.Y., S.I., and M. Matsuoka wrote the manuscript.

Figure legends

Fig. 1. Significance and dynamics of Tax expression at a single-cell level in MT-1 cells

(*A*) Single-cell qPCR for *tax* and *HBZ* expression in MT-1 cells (n=71).

(*B*) Left: Efficiency of Tax knockdown by shRNA. Right: The percentage of cells those are dead after Tax knockdown. MT-1 cells were infected with a bicistronic lentivirus vector expressing GFP and shRNA (shTax1, shTax4, or shNC). Cells were stained with LIVE/DEAD reagent and the ratio of dead cells to shRNA-transduced (GFP+) cells was measured by flow cytometry.

(*C*) AnnexinV/7-AAD double staining in shRNA-transduced MT-1 cells at day 15 post-transduction.

(*D*) Left: Schematic depicting the concept behind the GFP competition assay. The ratio

of GFP+ cells changes over time depending on the effect of shRNA upon transduced cells. Right: Serial measurements of the percentage of GFP+ cells among the whole MT-1 cell population after shNC or shTax4 transduction. For (B) and (D), error bars show SD for three experiments.

(*E*) Upper: Scheme of Tax reporter cassette that expresses d2EGFP. MT1GFP is a stable subline of MT-1 transduced with this cassette. Lower: Intracellular Tax staining of MT1GFP.

(*F*) Live cell imaging of Tax expression in MT1GFP cells. This montage of time-lapse images shows changes in d2EGFP expression.

(*G*) Expression pattern of Tax in MT1GFP. Cells were categorized based on their pattern of d2EGFP expression during the observation period. Transient: cell has single spontaneous episode. Fluctuating: cell has multiple episodes. Continuous: cell continuously expressed d2EGFP until the end of observation period.

(*H*) Single-cell dynamics of d2EGFP expression for 87 cells with transient expression. Mean ± SD is shown.

Fig. 2. Differences in gene expression between Tax-expressing and non-expressing MT1GFP cells

(*A*) RNA-Seq analysis comparing Tax+ and Tax- MT1GFP and KK1GFP cells. Read coverage plots for two Tax associated genes: TNFRSF9 (upper panel) and TRAF1 (lower panel).

(*B*) Ingenuity Pathway Analysis for dysregulated genes. Inclusion criteria are FPKM ≥ 1 and ≥ 2 fold expression change. Upper panel: The number of dysregulated genes

common to both MT1GFP and KK1GFP. Lower panel: Top-scoring activated or inactivated upstream regulators.

(*C*) Heatmap of single-cell gene expression for viral and cellular genes in single MT1GFP cells. Two experiments were performed; either d2EGFP+ enriched or un-enriched (whole) populations were subjected to single-cell analysis. Ninety-four cells from each experiment were analyzed.

(*D*) Violin plots comparing expression of viral and apoptosis-related genes in the Tax+ vs. Tax- populations. Based on the Apoptotic process gene list (GO:0006915, Gene Ontology database), 9 apoptosis-related genes, whose expression levels were significantly correlated with that of *tax* (r value > 0.5 by Pearson's correlation test), were identified.

(*E*) 3D PCA plot showing single-cell data clustering based on expression of apoptosis related genes from (*D*).

Fig. 3. Induction of Tax in MT1GFP cells by cytotoxic stresses and HIV-reactivating reagents

(*A*) Effect of cellular overgrowth on Tax expression. On day 4, cells were either passaged or allowed to overgrow. Upper panel shows cell count and bottom panel shows percentage of cells that are d2EGFP+ (Tax+).

(*B*) Effect of cellular overgrowth on cell viability in d2EGFP+ (Tax+) and d2EGFP- (Tax-) cells. Viability is measured by LIVE/DEAD staining.

(*C*) Induction of Tax expression by H_2O_2 treatment. Left: Percentage of cells that are

d2EGFP+ (Tax+); right: viability (LIVE/DEAD staining) of d2EGFP+ (Tax+) and d2EGFP- (Tax-) cells after 48 hours.

(*D*) Induction of Tax expression by cisplatin treatment is reversed by the antioxidant N-acetylcysteine (NAC).

(*E* and *F*) Effect of HIV-reactivating reagents on Tax expression in MT1GFP cells. (*E*) The effect of HIV-reactivating reagents on Tax expression. MT1GFP cells were treated for 18 hours with the indicated drugs that have been previously reported to reactivate HIV expression. (*F*) The effect of a combination of JQ-1 and PMA/I on Tax expression. The percentage of cells that are d2EGFP+ was measured by flow cytometry. Each figure (*A*-*F*) is a representative for two independent experiments. Error bars show SD for three replicates in one experiment.

Fig. 4. Cell cycle transition in d2EGFP+ (Tax+) and d2EGFP- (Tax-) cells

(*A*) Scheme for analysis of cell cycle transition without cell synchronization.

(*B* and *C*) Cell cycle phase distributions for EdU+ cells among d2EGFP+ (Tax+) and d2EGFP- (Tax-) cells at 2 and 6 hours. (*B*) Upper: gating strategy for EdU+ cells; lower: linear flow cytometric analysis of cell cycle with DAPI staining. Distribution of EdU+ cells is shown in red, and that of whole cells is blue.

(*C*) Bar charts showing the percentage of EdU+ cells in the indicated phase.

Fig. 5. Agent-based simulations of cell population dynamics under normal and Tax-knockdown conditions

(*A*) Scheme for modeling of cell population dynamics as a birth-death (Poisson) process with a stage transition. The ABM simulates the population dynamics of MT-1 cells transduced with shNC or shTax4. Top panel: normal conditions; bottom panel: Tax knockdown conditions.

(*B* and *C*) The time-course of the absolute number (*B*) and fraction (*C*) of shNC (control) and shTax4 (Tax knockdown) cells in a simulation resembling the experiment in Fig. 1*D*. In this simulated experiment, some MT-1 cells are shRNA+ and some are shRNA-. The simulated population dynamics of the shRNA+ cells in the mixed population are shown. The blue and green lines give the best-fit solutions for the agent-based simulations based on Tax-period sampling from experimental values (i.e., black bars in Fig. S3*C*) and Tax-interval sampling from simulated values (i.e., Fig. S3*G*). All data were fitted simultaneously.

(*D*) The simulated dynamics of the frequency of Tax-expressing cells in the normal MT-1 cell population. A steady state is reached in which about 3% of cells are Tax+.

(*E* and *F*) The simulated dynamics of subpopulations of cells transduced with shNC (*E*, blue lines) or shTax4 (*F*, green lines).

(*G*) The survival probabilities for subpopulations of shNC cells which have expressed Tax at least once and have never expressed Tax are described in solid and dashed lines, respectively.

Fig. 6. Proposed model of the dynamics and significance of Tax expression

Left: under normal conditions; right: under Tax-knockdown conditions.

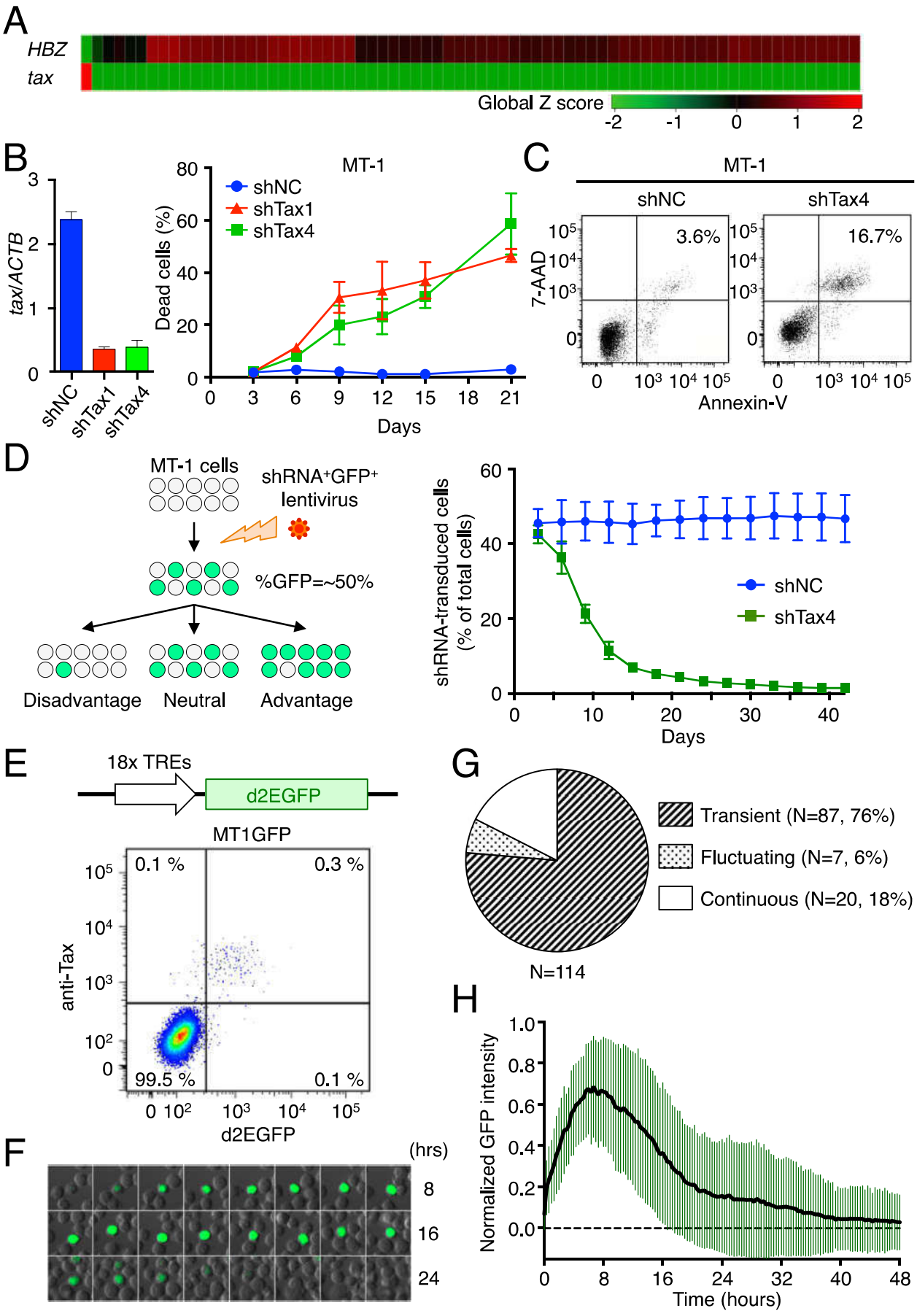

A
HBZ
tax
Global Z score
-2
-1
0
1
2
B
tax/ACTB
shNC
shTax1
shTax4
MT-1
Dead cells (%)
shNC
shTax1
shTax4
Days
C
MT-1
shNC
shTax4
7-AAD
3.6%
16.7%
Annexin-V
D
MT-1 cells
shRNA+GFP+ lentivirus
%GFP=~50%
Disadvantage
Neutral
Advantage
shRNA-transduced cells (% of total cells)
shNC
shTax4
Days
E
18x TREs
d2EGFP
MT1GFP
0.1 %
0.3 %
99.5 %
0.1 %
anti-Tax
d2EGFP
F
(hrs)
8
16
24
G
Transient (N=87, 76%)
Fluctuating (N=7, 6%)
Continuous (N=20, 18%)
N=114
H
Normalized GFP intensity
Time (hours)

Figure 1

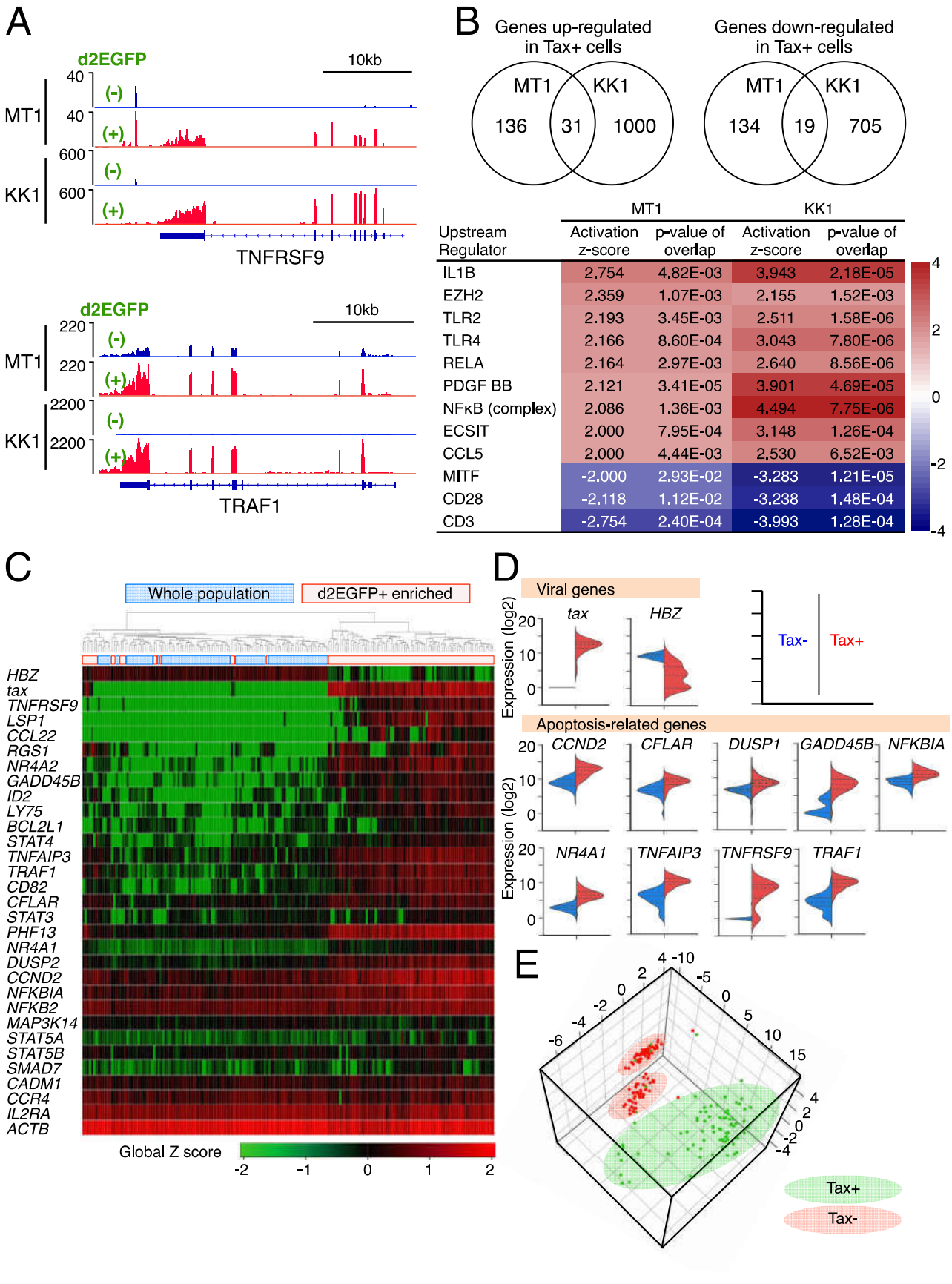

A
d2EGFP
10kb
MT1
KK1
(-)
(+)
40
600
TNFRSF9
d2EGFP
10kb
220
2200
TRAF1
B
Genes up-regulated in Tax+ cells
MT1
KK1
136
31
1000
Genes down-regulated in Tax+ cells
134
19
705
Upstream Regulator
Activation z-score
p-value of overlap
IL1B 2.754 4.82E-03 3.943 2.18E-05
EZH2 2.359 1.07E-03 2.155 1.52E-03
TLR2 2.193 3.45E-03 2.511 1.58E-06
TLR4 2.166 8.60E-04 3.043 7.80E-06
RELA 2.164 2.97E-03 2.640 8.56E-06
PDGF BB 2.121 3.41E-05 3.901 4.69E-05
NFκB (complex) 2.086 1.36E-03 4.494 7.75E-06
ECSIT 2.000 7.95E-04 3.148 1.26E-04
CCL5 2.000 4.44E-03 2.530 6.52E-03
MITF -2.000 2.93E-02 -3.283 1.21E-05
CD28 -2.118 1.12E-02 -3.238 1.48E-04
CD3 -2.754 2.40E-04 -3.993 1.28E-04
4
2
0
-2
-4
C
Whole population
d2EGFP+ enriched
HBZ
tax
TNFRSF9
LSP1
CCL22
RGS1
NR4A2
GADD45B
ID2
LY75
BCL2L1
STAT4
TNFAIP3
TRAF1
CD82
CFLAR
STAT3
PHF13
NR4A1
DUSP2
CCND2
NFKBIA
NFKB2
MAP3K14
STAT5A
STAT5B
SMAD7
CADM1
CCR4
IL2RA
ACTB
Global Z score
-2
-1
0
1
2
D
Viral genes
tax
HBZ
Expression (log2)
Tax-
Tax+
Apoptosis-related genes
CCND2
CFLAR
DUSP1
GADD45B
NFKBIA
NR4A1
TNFAIP3
TNFRSF9
TRAF1
E
Tax+
Tax-

Figure 2

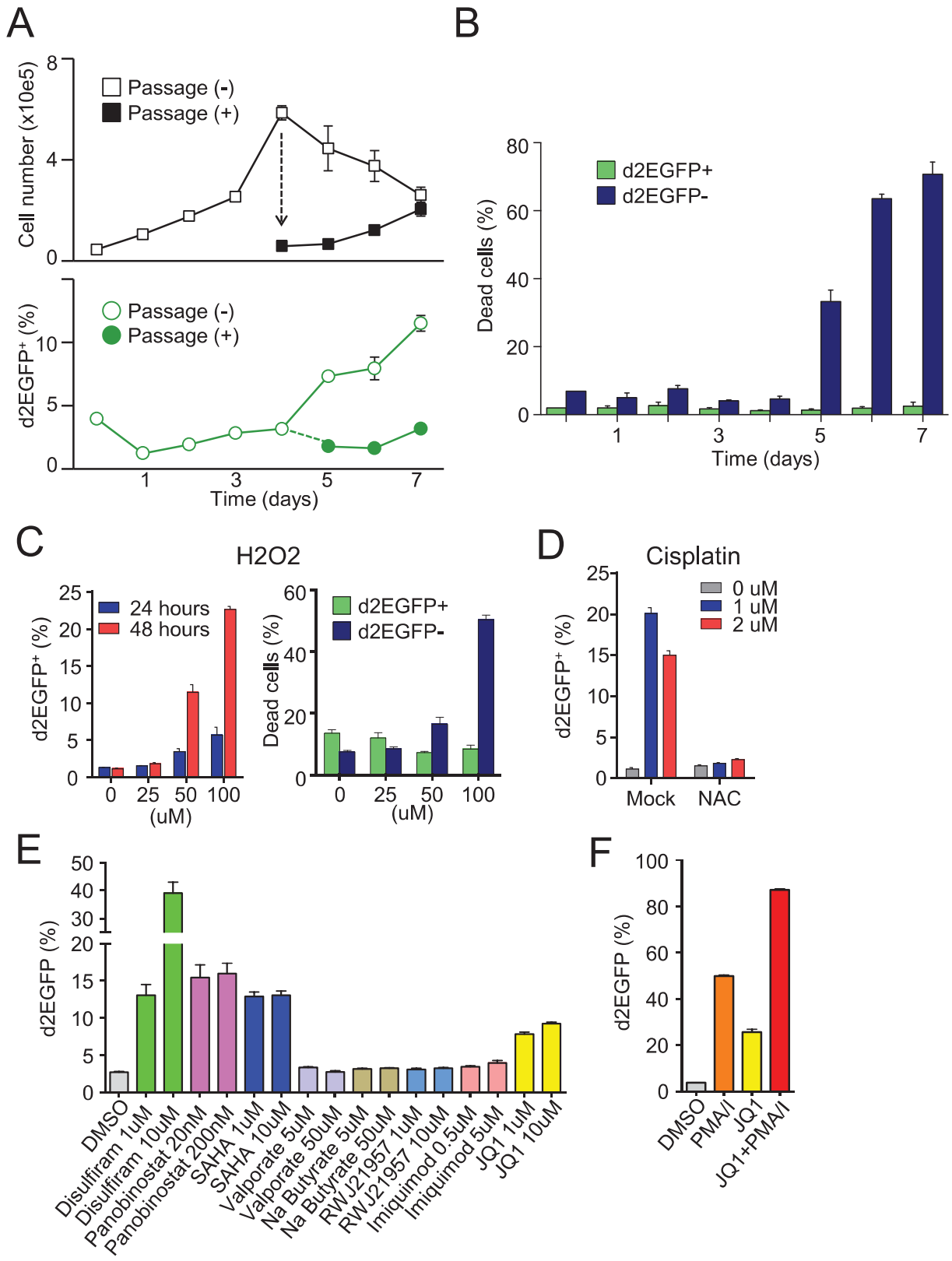

A
Cell number (x10e5)
Passage (-)
Passage (+)
d2EGFP+ (%)
Passage (-)
Passage (+)
Time (days)
B
Dead cells (%)
d2EGFP+
d2EGFP-
Time (days)
C
H2O2
d2EGFP+ (%)
24 hours
48 hours
(uM)
Dead cells (%)
d2EGFP+
d2EGFP-
(uM)
D
Cisplatin
d2EGFP+ (%)
0 uM
1 uM
2 uM
Mock
NAC
E
d2EGFP (%)
DMSO
Disulfiram 1uM
Disulfiram 10uM
Panobinostat 20nM
Panobinostat 200nM
SAHA 1uM
SAHA 10uM
Valporate 5uM
Valporate 50uM
Na Butyrate 5uM
Na Butyrate 50uM
RWJ21957 1uM
RWJ21957 10uM
Imiquimod 0.5uM
Imiquimod 5uM
JQ1 1uM
JQ1 10uM
F
d2EGFP (%)
DMSO
PMA/I
JQ1
JQ1+PMA/I

Figure 3

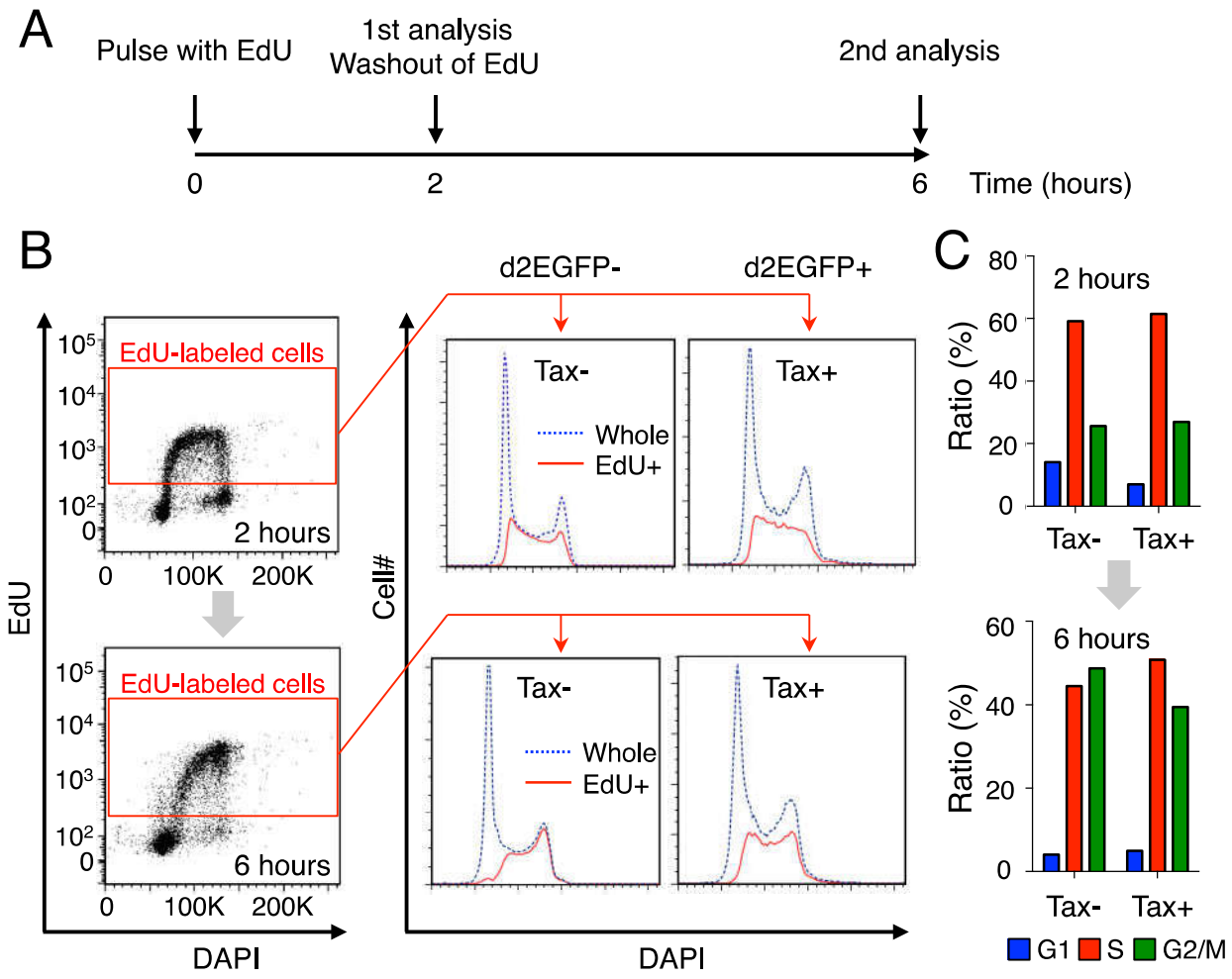

A
Pulse with EdU
1st analysis
Washout of EdU
2nd analysis
0
2
6
Time (hours)
B
d2EGFP-
d2EGFP+
EdU-labeled cells
2 hours
6 hours
EdU
DAPI
Cell#
Tax-
Tax+
Whole
EdU+
C
2 hours
6 hours
Ratio (%)
Tax-
Tax+
G1
S
G2/M

Figure 4

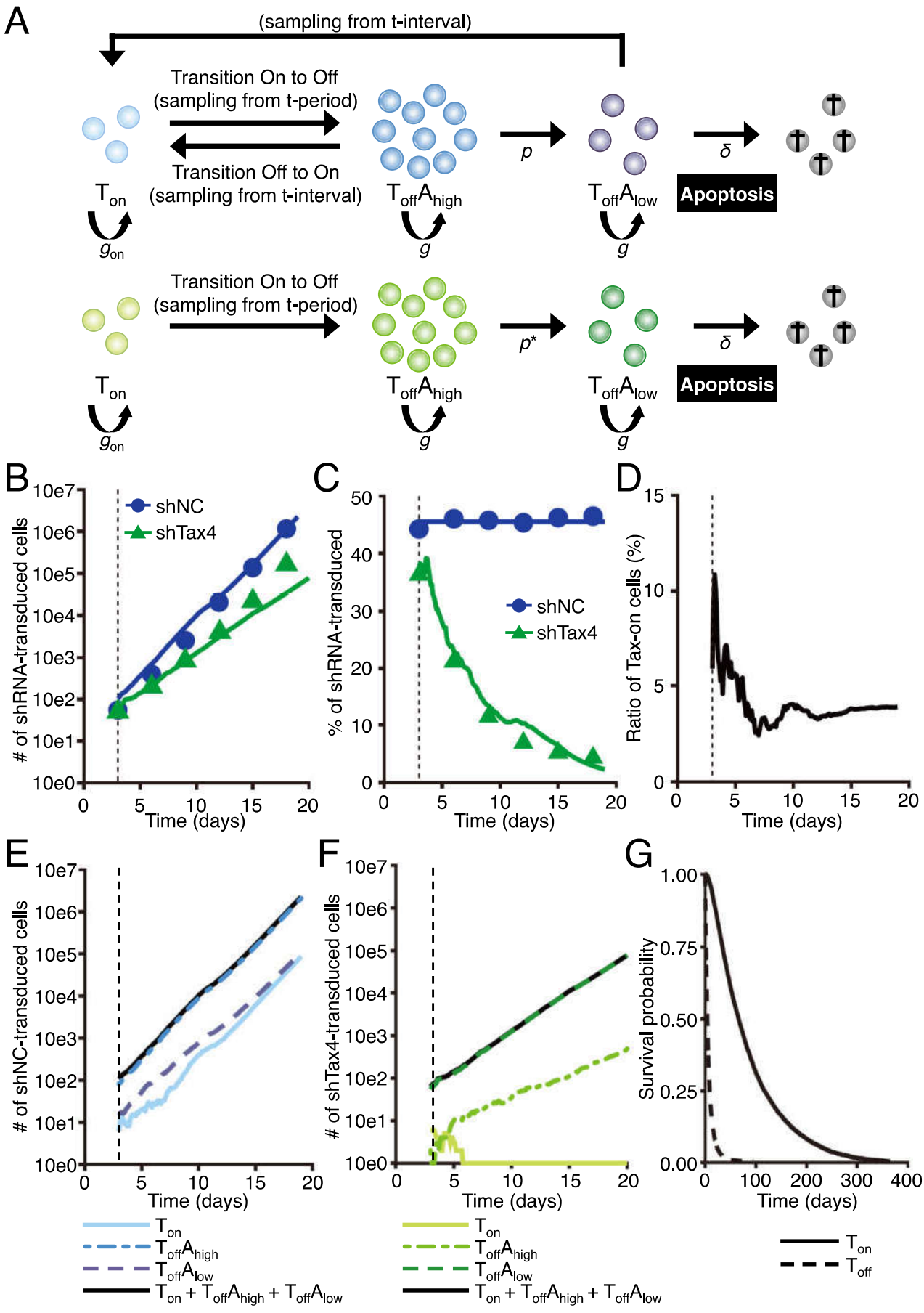

A
(sampling from t-interval)
Transition On to Off
(sampling from t-period)
Transition Off to On
(sampling from t-interval)
T_on
g_on
T_offA_high
g
p
T_offA_low
g
δ
Apoptosis
Transition On to Off
(sampling from t-period)
T_on
g_on
T_offA_high
g
p*
T_offA_low
g
δ
Apoptosis
B
# of shRNA-transduced cells
shNC
shTax4
Time (days)
C
% of shRNA-transduced
shNC
shTax4
Time (days)
D
Ratio of Tax-on cells (%)
Time (days)
E
# of shNC-transduced cells
Time (days)
T_on
T_offA_high
T_offA_low
T_on + T_offA_high + T_offA_low
F
# of shTax4-transduced cells
Time (days)
T_on
T_offA_high
T_offA_low
T_on + T_offA_high + T_offA_low
G
Survival probability
Time (days)
T_on
T_off

Figure 5

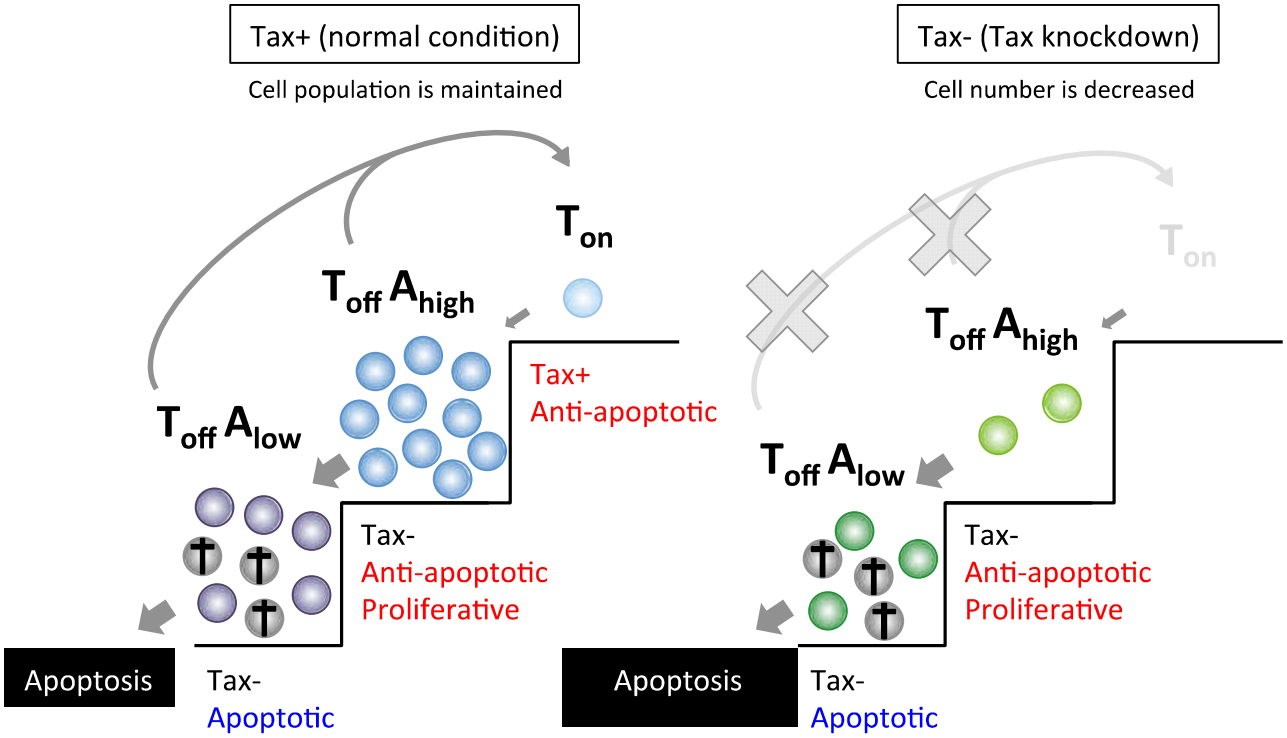

Tax+ (normal condition)
Cell population is maintained
Tax- (Tax knockdown)
Cell number is decreased
T on
T off A high
T off A low
Tax+
Anti-apoptotic
Tax-
Anti-apoptotic
Proliferative
Apoptosis
Tax-
Apoptotic
T on
T off A high
T off A low
Tax-
Anti-apoptotic
Proliferative
Apoptosis
Tax-
Apoptotic

Figure 6

Supplementary Material

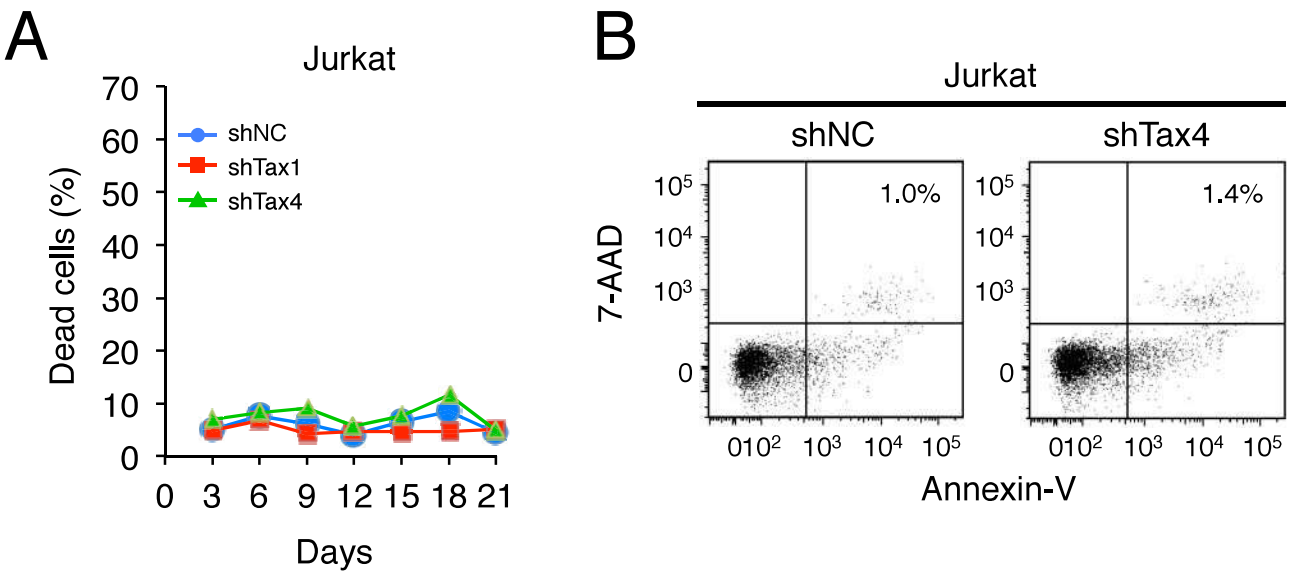

A
Jurkat
Dead cells (%)
70
60
50
40
30
20
10
0
shNC
shTax1
shTax4
0 3 6 9 12 15 18 21
Days
B
Jurkat
shNC
shTax4
7-AAD
10^5
10^4
10^3
0
1.0%
1.4%
$0\,10^2$ 10^3 10^4 10^5
Annexin-V

Fig. S1. Effects of shRNAs against Tax on HTLV-1-uninfected cells

(*A*) Tax knockdown does not cause apoptosis in uninfected cells. Jurkat cells were transduced with anti-Tax shRNA or control non-target shRNA. The percentage of dead cells in the transduced (GFP+) population was measured by flow cytometry using LIVE/DEAD reagent.

(*B*) AnnexinV/7-AAD double staining in shRNA-transduced Jurkat cells at day 15 post-transduction.

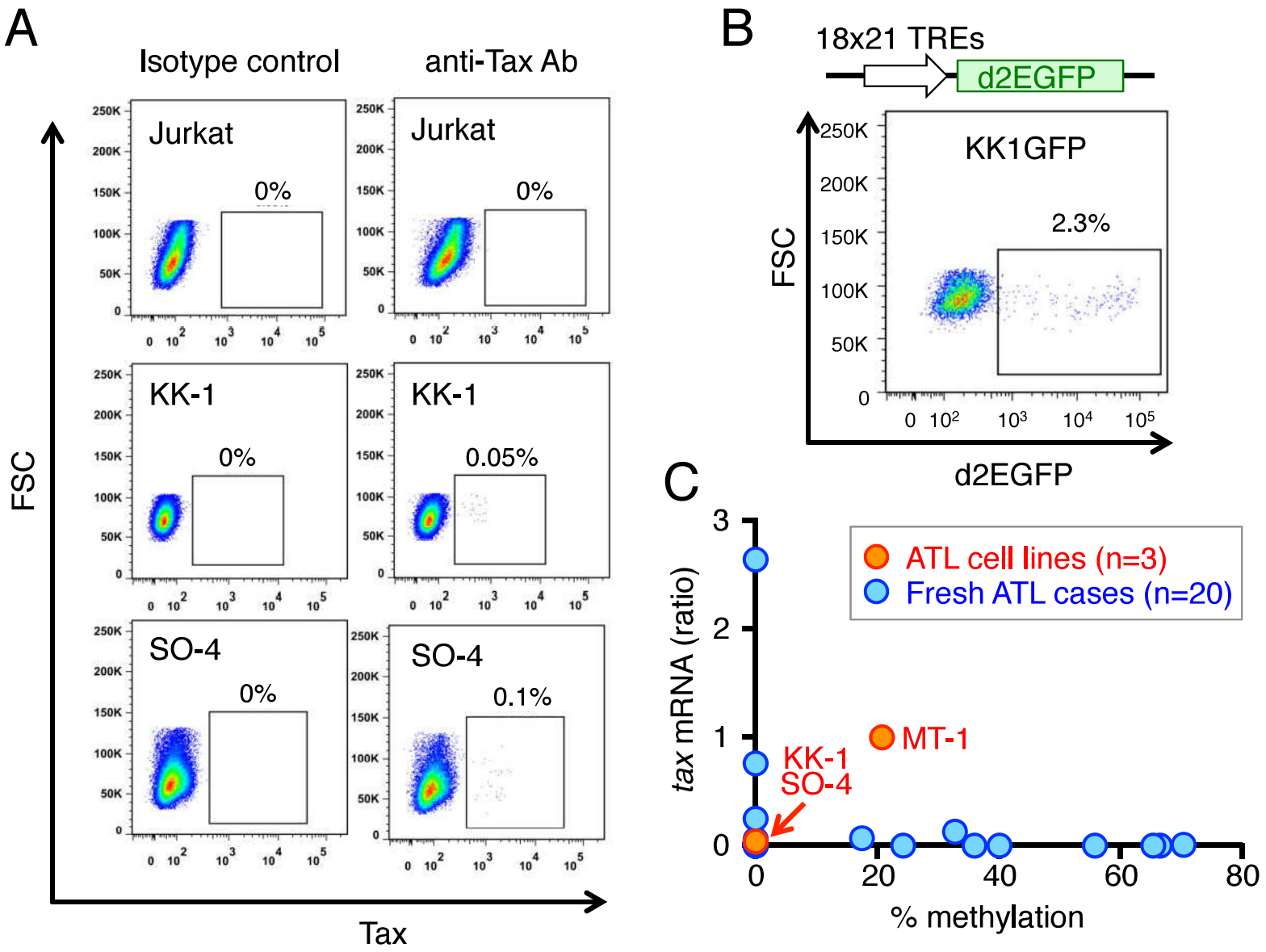

A
Isotype control
anti-Tax Ab
Jurkat
0%
Jurkat
0%
KK-1
0%
KK-1
0.05%
SO-4
0%
SO-4
0.1%
FSC
Tax
B
18x21 TREs
d2EGFP
KK1GFP
2.3%
FSC
d2EGFP
C
ATL cell lines (n=3)
Fresh ATL cases (n=20)
tax mRNA (ratio)
KK-1
SO-4
MT-1
% methylation

Fig. S2. Tax expression in several ATL cell lines

(*A*) Intracellular Tax staining in ATL cells lines KK1and SO4, and in an HTLV-1 negative T-cell line, Jurkat.

(*B*) Generation of a Tax reporter subline of KK1 (KK1GFP). Upper panel: Scheme of the reporter cassette expressing d2EGFP under the control of tandem repeats of TRE. Lower: Expression of d2EGFP in a representative subclone of KK1GFP (#11).

(*C*) Correlation between levels of *tax* mRNA and DNA methylation of 5'LTR in fresh ATL cells and ATL cell lines. Fresh ATL cells from aggressive type ATL (n=20) and three ATL cell lines (MT-1, KK-1, and SO-4) were subjected to the analysis.

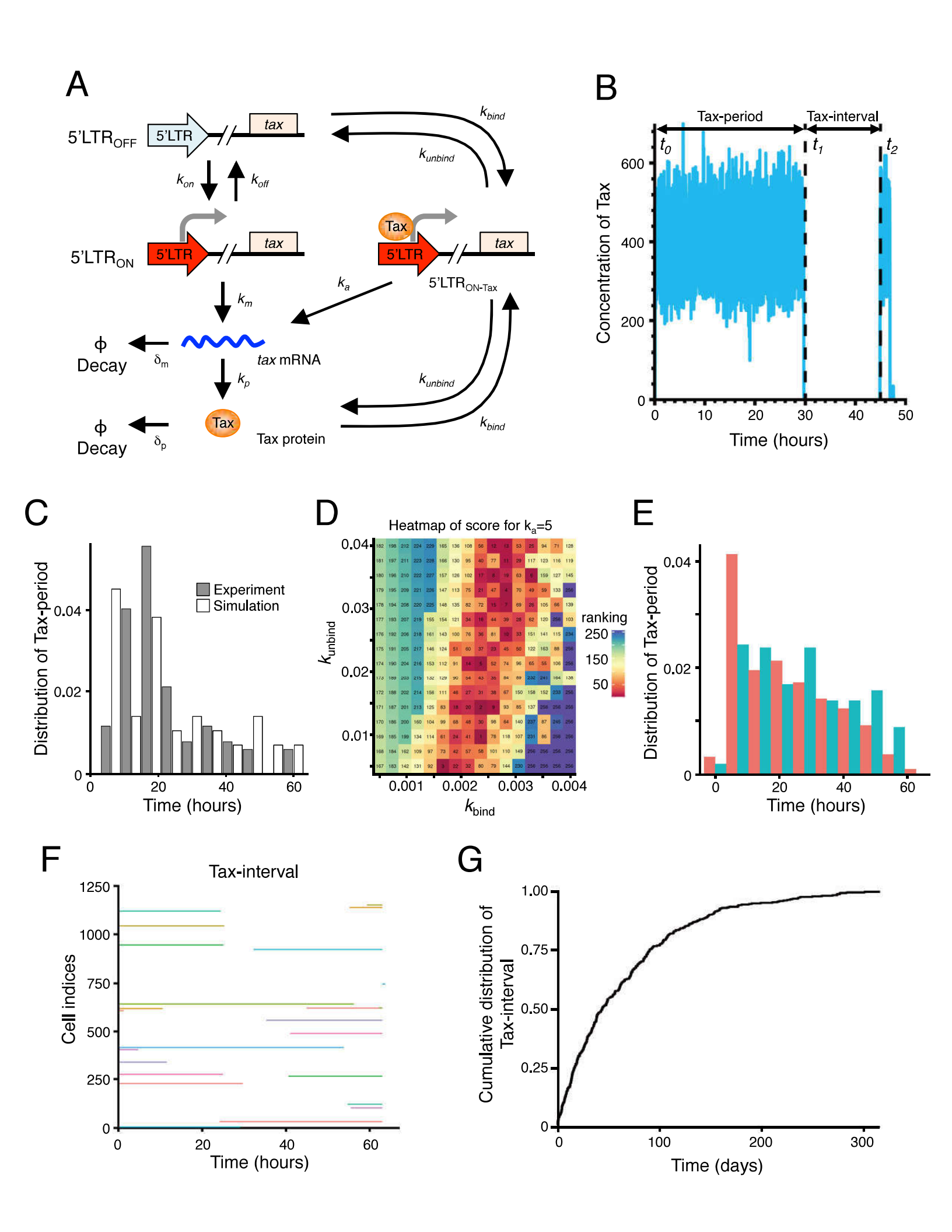

A
5'LTR_OFF
5'LTR
tax
k_bind
k_unbind
k_on
k_off
5'LTR_ON
Tax
5'LTR_ON-Tax
k_a
k_m
ϕ Decay
δ_m
tax mRNA
k_p
ϕ Decay
δ_p
Tax protein
B
Tax-period
Tax-interval
t_0
t_1
t_2
Concentration of Tax
Time (hours)
C
Experiment
Simulation
Distribution of Tax-period
Time (hours)
D
Heatmap of score for k_a=5
k_unbind
k_bind
ranking
E
Distribution of Tax-period
Time (hours)
F
Tax-interval
Cell indices
Time (hours)
G
Cumulative distribution of Tax-interval
Time (days)

Fig. S3. Computational simulation of intracellular Tax expression dynamics

(*A*) Scheme for the mathematical modeling of the Tax expression process.

(*B*) A sample path of the stochastic simulation of the Tax expression process. The *Tax-period* is defined as the time period of sustained Tax expression between the start point (the first time to reach the burst threshold of 200 copies from no expression) and the end point (the time to return to no expression); i.e., $t_{\text{period}} = t_1 - t_0$. The *Tax-interval* is defined as the time from the end of one Tax-period to the beginning of the next; i.e. the time during which the Tax concentration never reaches the burst threshold (i.e., $t_{\text{interval}} = t_2 - t_1$).

(*C*) The experimental and simulation-predicted distributions of Tax-period durations are shown.

(*D*) The ranking of feasible parameter sets determined by computing the closeness of simulated and experimentally determined Tax-periods. For each choice of the two parameters k_{bind} and k_{unbind}, a distribution of Tax-periods was generated by a single long-run stochastic simulation (up to 1.0×10^8 hours). The closeness between simulated and experimental Tax-periods was then numerically computed by the discretized Kullback-Leibler divergence (R package entropy). The ranking is determined by the closeness: the box with the minimum KL divergence value is ranked 1. If no tax-periods are found, the maximum KL divergence value is assigned (boxes with Ranking 256).

(*E*) Comparison of histograms for the distribution of Tax-periods between the collection of short-run (up to 65 hours) and long-run (65000 hours) stochastic simulations. A total

of 362 Tax-periods from 10,000 short-run simulations are used (red bars), while 407 Tax-periods were obtained from a single long-run simulation (blue bars).

(*F*) Plots of Tax-periods generated from short-run stochastic simulations. Twenty four samples that exhibit a Tax-interval are mapped on the time axis.

(*G*) The simulation-predicted cumulative distribution of Tax-intervals is shown. 50% of Tax-intervals are longer than about 40 days.

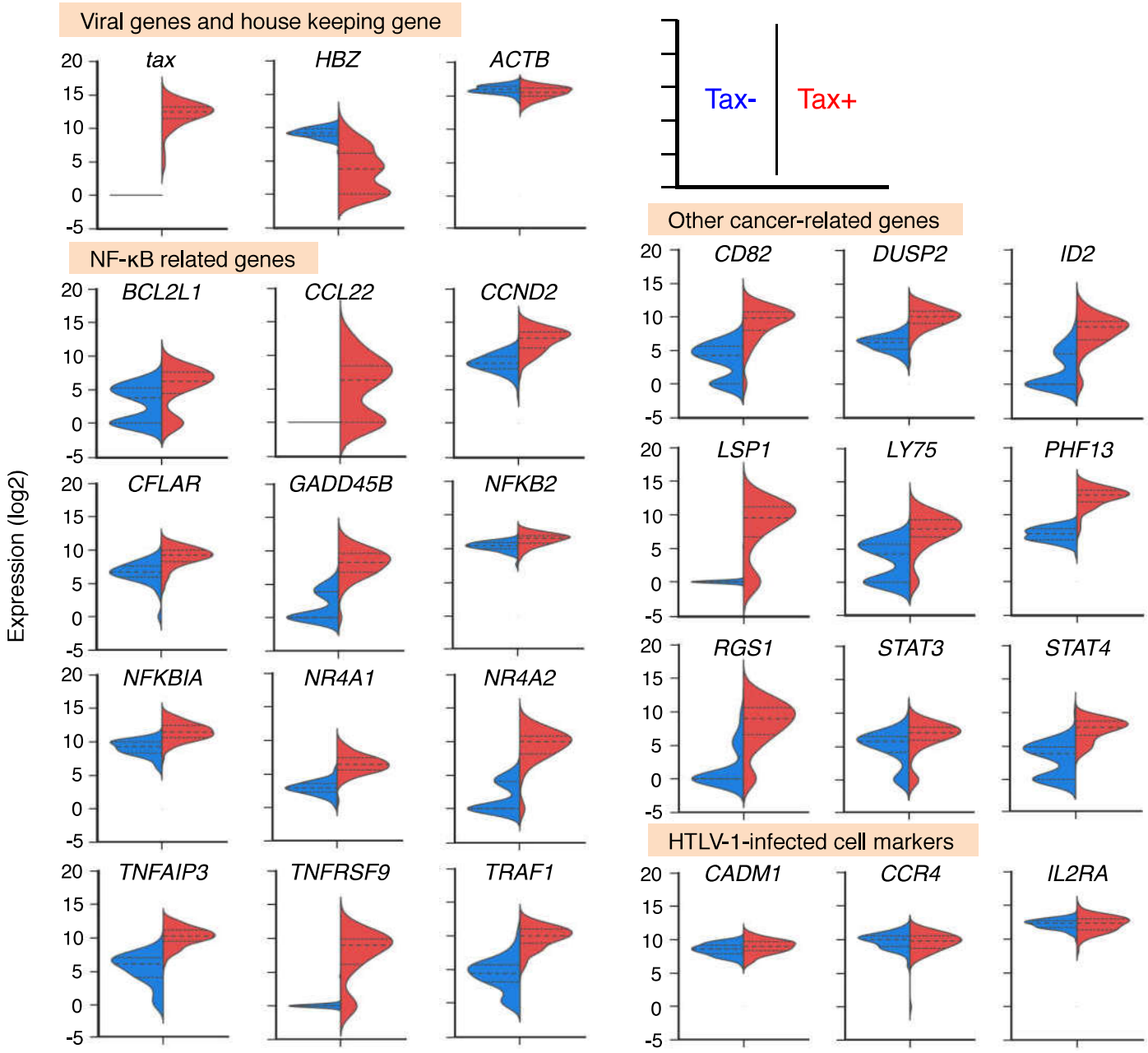

Viral genes and house keeping gene
tax
HBZ
ACTB
Tax-
Tax+
NF-κB related genes
BCL2L1
CCL22
CCND2
CFLAR
GADD45B
NFKB2
NFKBIA
NR4A1
NR4A2
TNFAIP3
TNFRSF9
TRAF1
Other cancer-related genes
CD82
DUSP2
ID2
LSP1
LY75
PHF13
RGS1
STAT3
STAT4
HTLV-1-infected cell markers
CADM1
CCR4
IL2RA
Expression (log2)
20
15
10
5
0
-5

Fig. S4. Violin plots of genes expressed differentially in Tax+ and Tax- MT1GFP cells

Results of single cell qPCR for the selected genes are shown as violin plots. Inner lines represent the population's distribution percentiles: upper: 75th; middle: 50th; lower: 25th.

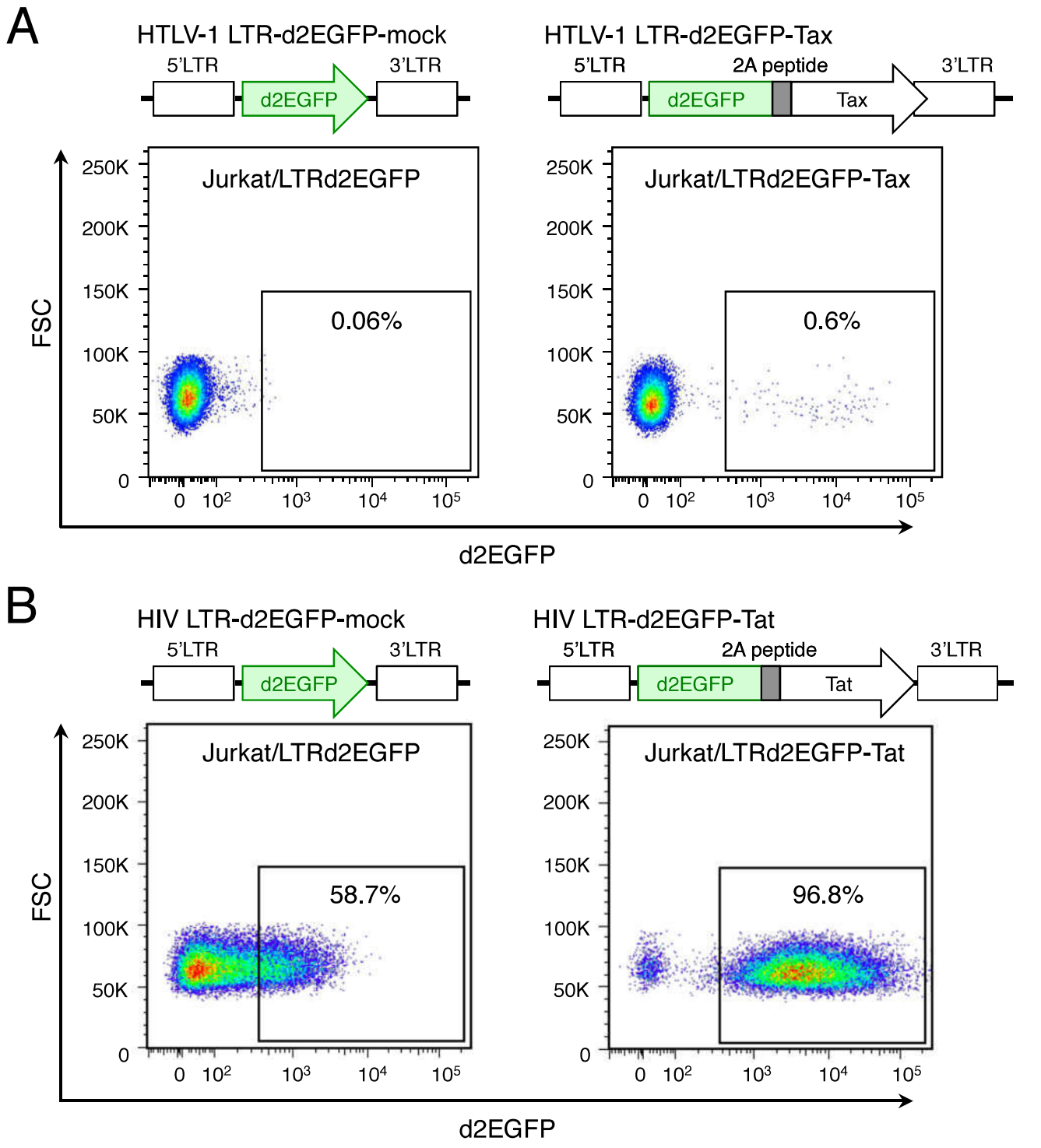

A
HTLV-1 LTR-d2EGFP-mock
5'LTR
d2EGFP
3'LTR
HTLV-1 LTR-d2EGFP-Tax
5'LTR
2A peptide
d2EGFP
Tax
3'LTR
Jurkat/LTRd2EGFP
0.06%
Jurkat/LTRd2EGFP-Tax
0.6%
FSC
d2EGFP
B
HIV LTR-d2EGFP-mock
HIV LTR-d2EGFP-Tat
Tat
Jurkat/LTRd2EGFP
58.7%
Jurkat/LTRd2EGFP-Tat
96.8%
C
Untreated
PMA/I + JQ1
d2EGFP (%)
#2
#3
#8
#20
#22
#28
Clone ID
D

Fig. S5. Latent phenotype established by HTLV-1-LTR-Tax minimal circuit

(*A* and *B*) d2EGFP expression under the control of the LTR promoter from HTLV-1 or HIV-1 in presence or absence of transactivator (Tax or Tat respectively) in stably transfected Jurkat cells. (*A*) d2EGFP expression from the HTLV-1 LTR: basal transcription (left) and transcription in the presence of Tax positive feedback (right). (*B*) d2EGFP expression from the HIV-1 LTR: basal transcription (left) and transcription in the presence of Tat positive feedback (right).

(*C*) Reactivation of LTR-Tax latency. Several Jurkat subclones stably transfected with HTLV-1-LTR-d2EGFP-2A-Tax (Jurkat/LTRd2EGFP-Tax cells) were established. All subclones had a similar latent phenotype, and could be drastically reactivated by PMA/Ionomycin + JQ1 treatment.

(*D*) Transient expression of Tax and d2EGFP in Jurkat/LTRd2EGFP-Tax cells. White arrowheads indicate cells transiently expressing Tax.

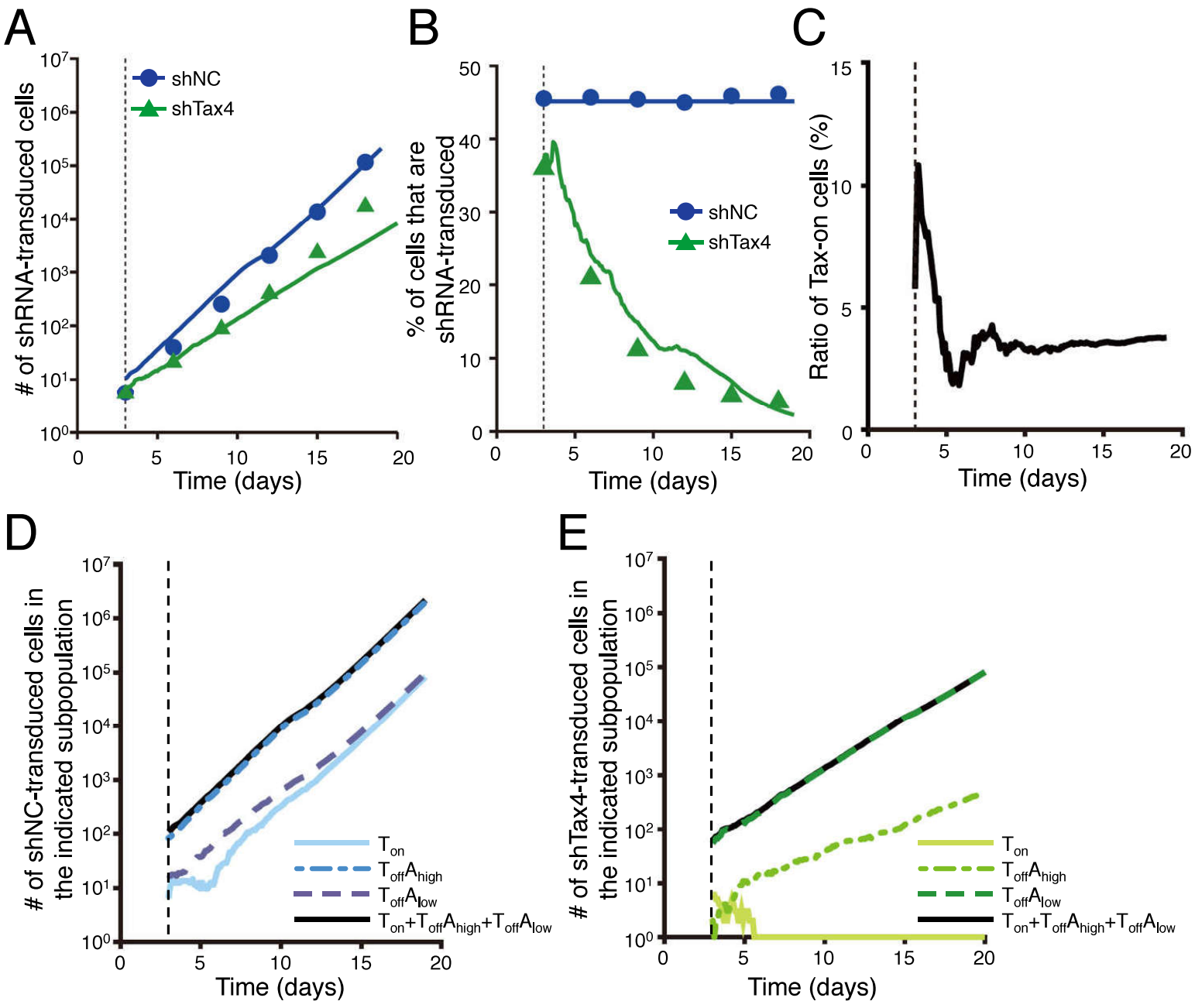

A
B
C
shNC
shTax4
# of shRNA-transduced cells
% of cells that are shRNA-transduced
Ratio of Tax-on cells (%)
Time (days)
D
E
# of shNC-transduced cells in the indicated subpopulation
# of shTax4-transduced cells in the indicated subpopulation
T_{on}
$T_{off}A_{high}$
$T_{off}A_{low}$
$T_{on}+T_{off}A_{high}+T_{off}A_{low}$

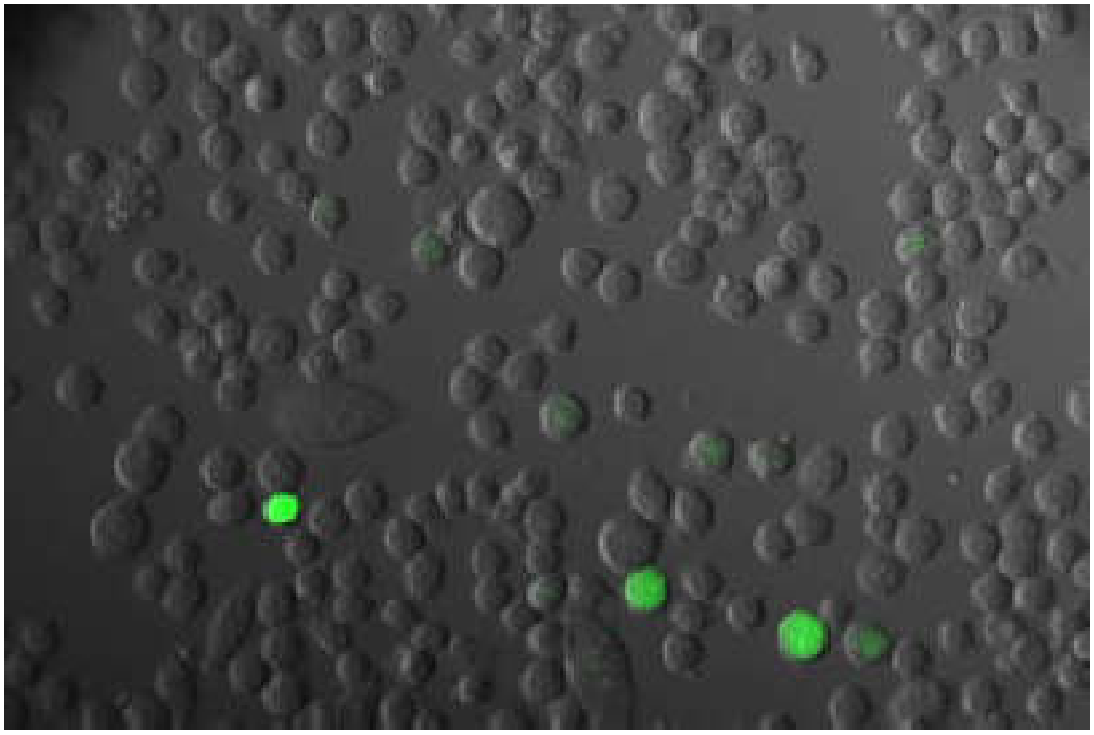
Fig. S6. An agent-based simulation using Tax-period sampling from previously simulated values

The simulation in this figure resembles the simulation in Fig. 5, except that Tax-period sampling is done from simulated values (i.e., the gray bars in Fig. S3*C*) and Tax-interval sampling from simulated values (i.e., Fig. S3*G*).

(*A* and *B*) Time-course of the number and fraction of shNC and shTax4 cells.

(*C*) The simulated dynamics of the frequency of Tax-positive cells in the normal MT-1 cell population.

(*D* and *E*) The simulated dynamics of subpopulations of cells transduced with shNC (*D*, blue lines) or shTax4 (*E*, green lines).



Movie S1. Time-lapse imaging of MT1GFP cells

Tax is transiently expressed in a small subpopulation of MT1GFP.

Table S1. Chemical reaction scheme and parameter values for stochastic simulation of the intracellular Tax expression model.

Reactions	Description	Rate and Value*
$5'LTR_{OFF} \leftrightarrow 5'LTR_{ON}$	Promoter toggling from active to inactive state	$^{\dagger}k_{ON} = 3.0 \times 10^{-6}$, $^{\ddagger}k_{OFF} = 1.0 \times 10^{-2}$
$5'LTR_{ON} \rightarrow 5'LTR_{ON} + mRNA$	Transcription of mRNA encoding Tax	$^{\ddagger}k_m = 0.1$
$mRNA \rightarrow mRNA + Tax$	Translation	$^{\ddagger}k_p = 10$
$5'LTR_{OFF} + Tax \leftrightarrow 5'LTR_{ON\text{-}Tax}$	Tax binding/unbinding to 5'LTR	$^{\dagger}k_{\text{bind}} = 2.5 \times 10^{-3}$, $^{\dagger}k_{\text{unbind}} = 1.8 \times 10^{-2}$
$5'LTR_{ON\text{-}Tax} \rightarrow 5'LTR_{ON\text{-}Tax} + mRNA$	Transactivated rate of transcription	$^{\ddagger}k_a = 5$
$mRNA \rightarrow \varphi$	mRNA decay	$^{\ddagger}\delta_m = 1$
$Tax \rightarrow \varphi$	Tax decay	$^{\ddagger}\delta_p = 0.125$

* Time in simulation is arbitrary units.

† These parameter values which are specific to HTLV-1 are estimated from experimental data by fitting.

‡ These parameter values are fixed as estimated in (Razooky et al., 2015).

Table S2. Parameter values for agent-based simulation of cell population dynamics under normal and Tax knockdown conditions.

Parameters	Description	Unit	Value
g_{on}	Growth rate of Tax-expressing cells	day^{-1}	†0.54
g	Growth rate of Tax-negative cells with or without anti-apoptotic gene expression	day^{-1}	†0.68
p	Decay rate of anti-apoptotic gene expression for $T_{\mathrm{off}}A_{\mathrm{high}}$ for shNC cells	day^{-1}	†0.02
p^{*}	Decay rate of anti-apoptotic gene expression for $T_{\mathrm{off}}A_{\mathrm{high}}$ for shTax4 cells	day^{-1}	†0.40
δ	Apoptosis rate for $T_{\mathrm{off}}A_{\mathrm{low}}$ cells	day^{-1}	†0.25

† These parameter values are estimated from experimental data by fitting.

Table S3. Primers used in the single-cell quantitative PCR.

	Pre-amplification primers		qPCR primers	
Gene	**Forward**	**Reverse**	**Forward**	**Reverse**
ACTB	CCTTCCAGCAGATGTGGATCAG	TCATCTTGTTTTCTGCGCAAGTTAG	AAGCAGGAGTATGACGAGTCC	CAAGAAAGGGTGTAACGCAACT
ATF3	GCTGTCTTCCTGTGTACCTCTA	AAATCCTGGAGTTGAGGCAAAG	CAGCAGCAGAGAACCATCAAG	CTGGTCCAAGACCCACTCTG
BCL2	GGATTGTGGCCTTCTTTGAGTT	CCAGCCTCCGTTATCCTGG	TGGGGTCATGTGTGTGGAGAG	GGTGCCGGTTCAGGTACTCA
BCL2A1	CGTCCAGAGTGCTACAAAATGT	CCGTCTTCAAACTCCTTTTCCA	GAAGTGGAAAAGAATCTGAAGTCATGC	TCACTTGGTTGAATAGTGTTCTGGC
BCL2L1	CCTAAGGCGGATTTGAATCTCTT	TATAATAGGGATGGGCTCAACCAG	CTCTCCCTTCAGAATCTTATCTTGG	ATTGTCCAAAACACCTGCTCAC
BCL3	CAGGATTTGCACAGAAGCACAT	GGAAATGGAAGATGTGGGTGG	CTCTTCTGAGCACAGATGTTCC	CATTCCGGACAGGGACTGG
BIRC5	GACCACCGCATCTCTACATTC	GCTCCTTGAAGCAGAAGAAACA	AAGAACTGGCCCTTCTTGGAG	AAGTCTGGCTCGTTCTCAGTG
CADM1	GAAGTACAGTATAAGCCTCAAGTGC	TTCATCATCGACTCTCACCCAA	CAGATGACTTATCCTCTACAAGGCT	GTTACCATCACAGGCTGGGG
CASP4	TGGGCTCTATCTTCATCACACA	TCATGGACAGTCGTTCTATGGT	GAAATATTCTTGGTGCTGCCAC	GCATTTGAGCTTTGGCCCTT
CASP8	CTTTGACCACGACCTTTGAAGA	CATGGGAGAGGATACAGCAGAT	TCATTTTGAGATCAAGCCCCAC	GTTACTGTGGTCCATGAGTTGG
CCL20	GTGCTTATCTAATTTGTGCCTCACT	TAGCTAAACACAGAAAACCTACAGC	CTAATTTGTGCCTCACTGGACT	TGTGATGCTTAAACAAAGCAAACT
CCL22	CTTCTACTGGACCTCAGACTCC	GGTCATCAGAGTAGGCTCTTCA	CGTGGTGTTGCTAACCTTCAG	TTGGCTCAGCTTATTGAGAATCA
CCL3	CTGCAACCAGTTCTCTGCATC	TTAGGAAGATGACACCGGGC	CTTGCTGCTGACACGCCG	GCTGCTCGTCTCAAAGTAGTCA
CCND2	GGAAGTTGAAGTGGAACCTGG	CAAACTTAAAGTCGGTGGCACA	TGACTTCATTGAGCACATCTTGC	AGAGCAATGAAGGTCTGAGCAT
CCR10	CTAGGGCTGCGAATCTAGAGG	GGGTTGCATCTCATTTCCATGT	GGTCGTGGGAAAGGGGAGTA	TTTAATGTGGCAAGGCACAGA
CCR4	AAAGCAAGCTGCTTCTGGTTG	TGCAAGGCTTGGGGATACTTT	GCCCAGACCTGCCTTGAG	TTGCTGTATATGCTTTCATCGAGGG
CCR6	GATACAGTCAACAAGCCTGACC	CAACCTGAGCTGCTTTATCCAA	AAATGGAGGTGGAAAGACAAGC	AATTGTCACCAACACCTGTTCT
CCR9	GGAAGTTCCAGACCTTCATGTG	TTTCTCCCTCCAAGTATGTGCT	GGTCAACAGCATGTACAAGATGAAC	GGCAATGGCAATGTACCTGT
CD48	ATGGAAGATCAAGCTGCAAGTG	TTTTGTCCCCATACCAGGTGTA	GTACCCAAGCCTGTCATCAAAA	TTTACAGACTCGCCAGGTATCA
CD55	CCCTAATCCGGGAGAAATACGA	GGACAGAGCTGCCTGAAATAAG	AATGGTCAGATTGATGTACCAGG	AATAATTTGTACCCTGTGTTACATGAGAAG
CD69	GAGCTCCAGCAAAGACTTTCAC	GGACTGGTGGCATCATTTTCTT	GACTTGACCTGAGATTAACTAGGGA	TCCACTCTCCGGATGCAAAG
CD82	GATGGGCTCAGCCTGTATCAA	GGAGGTTTGCAGGACAGAGAT	GTCACCAAATACTTTCTCTTCCTCTTC	GAAACTGCTCTTGTCGGCCA

CD99	TTAATCTTGCGATGTGCTTTGC	AATAATGGCCAGCCTTTCACAT	GCTGGGCGGATGATGTTTACTA	GGGCCTTGGAGAATGGGG
CDC25B	CTGGACAAGAGAGTCATCCTCA	TAGCCGCCTTTCAGGATATACA	TGTGAATTCTCATCTGAGCGTG	GTAGTAGAGGCTGGGGTAGTC
CDKN1A	ACCAGCATGACAGATTTCTACC	CACACAAACTGAGACTAAGGCA	CGGCTGATCTTCTCCAAGAGG	GAAGATGTAGAGCGGGCCTTT
CFLAR	CCCCTAGGAATCTGCCTGATAA	TACAGGCAAATTGGCCAAGAAT	ATTGGCAATGAGACAGAGCTTC	TGGGATATACCATGCATACTGAGA
COL1A1	AGACATGTTCAGCTTTGTGGAC	GTACCTGAGGCCGTTCTGTA	GGCTCCTGCTCCTCTTAGC	AGGTGATTGGTGGGATGTCTT
CREB3	GGCTAGTACTGACAGATGAGGA	GGCTCTCTTGAGCAGATCTTTTA	AGAGTCTATTGGAGAAGGAGGG	CGCACACGTTTCAGAATTTGTT
CXCL10	TGTCCACGTGTTGAGATCATTG	GATTTTGCTCCCCTCTGGTTTT	GGTGAGAAGAGATGTCTGAATCC	GAGATCTTTTAGACCTTTCCTTGCT
CXCL11	TGGGGTAAAAGCAGTGAAAGTG	ATAAGCCTTGCTTGCTTCGATT	AGCCTCCATAATGTACCCAAGT	GGATTTAGGCATCGTTGTCCTTT
CXCL8	AAACTTTCAGAGACAGCAGAGC	CAGAGCTGCAGAAATCAGGAAG	ACAAGCTTCTAGGACAAGAGCC	GCCAGCTTGGAAGTCATGTTTA
DUSP1	GAAATCCTGCCCTTTCTGTACC	AGGGATGCTCTTGTACTGGTAG	CAGTGCGTATCACGCTTCC	CCCTCAAAATGGTTGGGACAAT
DUSP2	GATCTTGCCCTACCTGTTCCT	TCCACAGGGATACTCTTGTAGC	CTGCAGTCACTCGTCAGACC	GAAAAGGCCCTCAAAGTGGTTG
EGR1	TAGGTCAGATGGAGGTTCTCAG	AGAACTTGGACATGGCTGTTTC	CAAGTCCTCCCTCTCTACTGGA	GCAGCTGAAGTCAAAGGGAATA
FZD6	CTTGGGTCTCTGATCATTGTCG	TTTTGCTTCCAACCCAGAAGAC	CCCATGTCCTTATCAGGCAAAA	TGCCAACAATTAATGTCATCAGGT
GADD45B	GAGTCGGCCAAGTTGATGAATG	ACGATGTTGATGTCGTTGTCAC	CTGCCTCTTGGCCATTGACGA	AGCAGAAGGACTGGATGAGCG
HIF1A	TACAGGATGCTTGCCAAAAGAG	GGAGAAAATCAAGTCGTGCTGA	CTGGGTTGAAACTCAAGCAACT	ACCACTCACAACGTAATTCACA
HIGD1A	ACAGGTGTTTCCCTTCCTTCAT	AGTATTTCCCCTGCTCTTCAGT	GGAAGATCAGGGATCAAAACTCA	CAACAATTGCTGCAAAACCCG
ICAM1	ATGTGCAAGAAGATAGCCAACC	GGTAAGGTTCTTGCCCACTG	GCTATTCAAACTGCCCTGATGG	GGAGGGGTGCCAGTTCCA
ID2	TCAAATGACAGCAAAGCACTGT	CCATTCAACTTGTCCTCCTTGT	GTGGCTGAATAAGCGGTGTTC	AAGTTCAGCACTTAAAAGATTCCGTG
IFIT2	TCAGCATTTATTGGTGGCAGAA	ATGGCATTTTAGTTGCCGTAGG	GGAAGATTTCTGAAGAGTGCAGC	CTGCTCTCCAAGGAATTCTTATTGT
IFIT3	GGCAGTATTTTCCTGTCAGCAT	AGGCTTCTGATGCTCTGTTTTC	TGAGCTTGAGGATGGTAGTGAG	TCTGTCTCAGTTCAGTTGCTCT
IKZF1	ATATTGGAGACCGGAGGGTAAC	TTACAGAAGGGATATGGGGCAA	GTTGGCCTGTTGATTACAGCTA	TGAGATCCAGGCTTTCCACAT
IL10RA	CTCCTAACCTCTGGAGAAGTGG	AGAAGATGATGACGTTGGTCAC	GAGTTCTGTGTCCAGGTGAAAC	GTGAAATACTGCCTGGTGAGG
IL13	TCATTGCTCTCACTTGCCTTG	GTTGATGCTCCATACCATGCTG	CCTGTGCCTCCCTCTACAGC	CATTGCAGAGCGGAGCCTTC
IL15	GAAGTGCTTTCTCTTGGAGTTACA	TCCTCACATTCTTTGCATCCAG	ATTTCACTTGAGTCCGGAGATG	GAAGACAAACTGTTGTTTGCTAGG
IL23A	TGATGTTCCCCATATCCAGTGT	AGAAGGCTCCCCTGTGAAAATA	CCCAAGGACTCAGGGACAAC	TCCGATCCTAGCAGCTTCTCAT
IL2RA	GGAAGACAAGGTGGACCCAG	TGAAAATCTGTTGTTGTGACGAGG	CAGCTCATATGCACAGGTGAAA	CAGGAAGTCTCACTCTCAGGAC

IMMT	GCTCGGGATGACTTTAAACGAG	TCAATACGACGATGTGCATGAG	AGCTGGACAGTATTACTCCAGAA	CATCAGTAGAGAGCTTGTCAGC
INSIG1	ACTGCATTAAACGTGTGGTGTT	ACAAGGCTCAGATTGGTTTTCA	GTGAGCACAATGTATTCATTTAATGGC	TGCACTTGAAGTTTAAGACCCA
JUNB	GCACTAAAATGGAACAGCCCTT	CTTTGAGACTCCGGTAGGGG	TACCACGACGACTCATACACAG	GCTCGGTTTCAGGAGTTTGTAG
LSP1	TATTGTGGCTGGAGACATGAGC	TCCACAAGCACCTTCTCATACT	AAGAAAAGCCTCTGGGAGCAGAA	CGGTGGCCACAAACTTATACCT
LY75	AAGAGTTGTCTGCAAAGTGCC	GAACTGATGAGAAACCCGCC	GGCCCTGATTACACAGCAATAG	GAAGAGGAACCAAATCAGTCCG
MAP3K14	GCAGCTGGAAATAGAATTATTCCTC	GAGCTTTGAGAGGCCTTTGATG	CCTGTCCCAGCCATTTTCTCT	GGTTCTTCTCACTGTCATCCGA
MT1X	CCACGCTTTTCATCTGTCCC	GTGCATTTGCACTCTTTGCATT	CTGCGTGTTTTCCTCTTGATCG	CAGGCACAGGAGCCAACAG
MT2A	AATGCACCTCCTGCAAGAAAAG	TGGAAGTCGCGTTCTTTACATC	TGGGCTGTGCCAAGTGTG	GGGCTGTCCCAGCATCAG
MYD88	CAATTCCTGGAGATGCCAACTT	GTCTCAGCTGGTCCTGTTCTAT	CAGCAGCTGGACATCACATTTC	GGCTGATAATCCAGCAAGTGG
NFKB1	CCGTGTAAACCAAAGCCCTAAA	GAACCAAGAAAGGAAGCCAAGT	ATTCCACTGCGTTGTCCAC	AAATCGAGAATGATTCAGGCGG
NFKB2	TGTTGCATATGCCTGACTTTGA	CTAGATGCAAGGCTGTTCGTC	CTGTATCCAGTACACCTGGCGG	CTCCACTTCAGCCCCACTGTC
NFKBIA	AATGCTCAGGAGCCCTGTAA	AGGTGAGCTGGTAGGGAGAAT	CCTTCACCTCGCAGTGGAC	CCCTGGTAGGTAACTCTGTTGA
NR4A1	ATAGCGCCGTGCTGTAAATAAG	GAAGGAGGGTAGGAAGAGAGAG	CAGGAAGAAAGAGCTTGAGGTG	CAGCAGGAGGCTGGAAAGG
NR4A2	GAATGAATGAAGAGAGACGCGG	TTACAGGCGTTTTCGAGGAAAT	CCTAAGGAGGAGATTGGACAGG	AGGTGGACAGTGTCGTAATTCA
PHF13	ATGATGAGGACATCATGGTGGA	ATTTCCGGATTTTCGCACAGG	CTCAGATGACGATTCCTGGGAC	TGGCACTCATTACACTCGATCA
PKMYT1	CCTCACCATCCTGGAAGTGG	CTCCAGCATCATGACAAGGAC	CATGCAACATGGAGCTGCCC	GAAGACAGACCGGCAGTGAA
PPP1R21	TGCGTTAAGAGCCAGGAATCTA	ACACAGAACTGTAGTTTGTCCG	TGACAGCTGTGTTTGAGAAGC	AAGGAGTCCATCTGGCTCTGTA
PRKACB	CTGGAGATACCAGCAACTTTGA	TCTCTCAACAGAGTGCAAACAC	TGACTATGAAGAAGAAGATATCCGTG	GTCATCTTGTTCCTCTTTAAAATTCACC
PYCARD	CCTAAGGGAGTCCCAGTCCTA	CAAGCTGGCTTTTCGTATATTGTG	CTGGTGGAGGACCTGGAG	TCAGGATGATTTGGTGGGATTG
RELA	GCTTGTAGGAAAGGACTGCC	TTGTTGTTGGTCTGGATGCG	GGATGGCTTCTATGAGGCTGAG	CCGCTTCTTCACACACTGGATT
RGS1	TGCAATGGTCTCAATCTCTGGA	GCTTTACAGGGCAAAAGATCAGA	TGCCAACCAAACTGGTCAAAAT	TTATAGTCTTCACAAGCCAGCC
SELPLG	CTCTAACATCTCCCATGGCTCT	AATGTCCACTTCCTGTTTGGTG	CTTCTCCTGGTCACTGGAGTC	CATGGCAACCAAAAGACAATGG
SMAD7	GCTTTCAGATTCCCAACTTCTTCT	CCATTCCCCTGAGGTAGATCAT	GGATCGGTCACACTGGTGC	AGAAGATATCCAGAGAGGGCTCC
SMIM4	TCCAGTGCTTAGTTCCGTCATA	ATCGTTCCTCCCAGGACAAAA	GGTCGAGCATGTTCACCAG	AGGAACCGGTAGATGCCAAATC
STAT1	GACCGCACCTTCAGTCTTTTC	GCAATTTCACCAACAGTCTCAAC	TCATTCAGAGCTCGTTTGTGG	AGTGAACTGGACCCCTGTCT
STAT3	CACTACTAAAGTCAGGTTGCTGG	CATCACTTTTGTGTTTGTGCCC	TTCCCTGAGTTGAATTATCAGCTT	AATGTTAAATTTCCGGGATCCTCT

STAT4	CTAATGCTTGGGCATCCATCAT	TCTGAGTTAAGACCACGACCAA	CAACGATTCCCAGAACTTGGTT	ATGAAAACTGCCAGCTCATCAC
STAT5A	CACAGAACCCTGACCATGTACT	AGAGGTGAAAAGACCGGCAG	GGATGGAGAATTCGACCTGGAT	CGGGAGTCAAGACTGTCCATT
STAT5B	CTACGTGTTTCCTGATCGGC	AGATGCGTTCACAAACTCAGG	TACTCCAAATACTACACACCAGTTC	GACCACTTGCTTGATCTGTGG
THEMIS2	AGAAGGTGGTCTGTGAGAACC	TGCTTGGAGCTCTGAGTTGT	CGAAGACCAGCCAGACCAT	CTCCTCCAGGGTTTCATAGCTC
TNFAIP3	GACACAGACTTGGTACTGAGGA	TCCAGTTCCGAGTATCATAGCA	TTCAGCACGCTCAAGGAAAC	AAGCCCCGTTTCAACAAATTCC
TNFRSF1B	CATGCCGGCTCAGAGAATACTA	AGTTCCAGAGCTGGGTGTATG	TGCAGCAAATGCTCGCCG	CTGTCCTCACAGGAGTCACAC
TNFRSF9	CAGGATGCAGCATGTGTGAA	GACTTTCCATCCAAAGAACAGTTTG	CAGGATTGTAAACAAGGTCAAGAAC	GACAGATGCCACGTTTCTGAT
TNFSF10	GGCATTCATTCCTGAGCAACTT	TGGACCATTTGTTTGTCGTTCT	CACTTGAGGAATGGTGAACTGG	TTTCCTCCTGAAATCGAAAGTATGT
TNFSF4	GTCTGTCAACTCCTTGATGGTG	AGGACACAGAATTCACCAGGAT	GCCTCTCTGACTTACAAAGACAA	AAGAATCAGTTCTCCGCCATTC
TNIP1	CTGGCCAAGTCCAAGATTGAAA	GCTGGATCTCCTTTTCCTGGTA	TGGAGGAGACCGACAAGGAG	CTCACGCTGTCGGGTGAGT
TP53	GCTCAGATAGCGATGGTCTGG	GTACAGTCAGAGCCAACCTCAG	CTCCTCAGCATCTTATCCGAGT	ATAGGGCACCACCACACTATG
TRAF1	CACTTTCCTGTGGAAGATCACC	TGCCATCTCCATTCAGGTACAG	AATGTCACCAGGCGGTGC	CATACTTGGCAGTGTAGAAGGC
UBE2C	ATAAGCTCTCGCTAGAGTTCCC	ATCATACAGGGCAGACCACTTT	CAGTGGCTACCCTTACAATGC	ATGTCCAGGCATATGTTACCCT
WIPI1	ATGGATTCACAATTCCTCTGCG	AGGTAGTTGGTAGCAGCCATAA	CCGAGACGGTACACATCTTCAA	ACATCTTTCCCATGTAGCCACT
tax	GATCCCAAAGAAAAAGACCACCA	CTCCAAACACGTAGACTGGGTA	GATCCCAAAGAAAAAGACCACCA	CTCCAAACACGTAGACTGGGTA
hbz	TAAACTTACCTAGACGGCGGAC	ACCAATTCCTCCACCAGCAG	CAGTTCAGGAGGCACCACAG	ACAGGCAAGCATCGAAACAG
Spike1	GTATTTGATTACCGGCAACGGA	ACCGTCGTCATTGAAGATATGC	GTGCATTCTCTGGAAGGTGAAC	CTCGGGTATCCCATTGCTGAT
Spike4	ATGCGTAGATAACATTCAGGCG	GTGGTGAAATTCGATCCTACCG	AAGAGCTGTACACCGTAGTGG	CCACATCTTTGGCAGCGAAT
Spike7	GAGAAAGGTAATGACCTGGTGC	TTTGGGTATTGAAGATCAGGCG	TCCCATATGCTGGAAGTCTGC	GATCGTCCTTACCACCAATACC

SI Methods

Plasmids

The *Piggy*Bac Transposon system (System Biosciences) was used in this study. A Tax-responsive element-driven reporter plasmid, PB-18X21-RFP, and a transposase-expression vector, PB-Tase, were kindly provided by Dr. Chou-Zen Giam (1). To obtain d2EGFP under the control of the TRE's, we generated PB-18X21-d2EGFP from PB-18x21-RFP and pd2EGFP (Clontech) by replacing the coding sequence of RFP with the fragment containing d2EGFP and the neomycin-resistance gene. The 2A peptide-based bicistronic construct expressing Tax (or Tat) and d2EGFP through the HTLV-1 5'LTR (or HIV 5'LTR) replaced the CMV promoter of PB-CMV-MCS-EF1-Puro (System Biosciences) to produce PB-HTLV-1 LTR-d2EGFP-Tax and PB-HIV LTR-d2EGFP-Tat, respectively. For knockdown experiments, lentivirus vectors expressing shRNAs were used. For shRNAs, a fragment targeting a Tax coding sequence (shTax1: gccttcctcaccaatgttc, or shTax4: ggcagatgacaatgaccatga) or a non-target shRNA (shNC: caacaagatgaagagcaccaa) was inserted into pLKO.1-EGFP (Sigma-Aldrich).

Transfection and lentivirus transduction

We stably transfected MT-1 cells with the reporter transposon PB-18X21-d2EGFP using the Neon Transfection System (Thermo Fisher Scientific), thus generating MT1GFP cells. $5x10^5$ MT-1 cells were electroporated with 1 μg of PB-18X21-d2EGFP and 0.3 μg of PB-Tase helper plasmid. 48 hours later, G418 (1000 μg/ml) was added for selection. Stable clones were isolated by limiting dilution and expanded in 96-well plates. For

lentivirus production, 3-5 x 10^6 of HEK293T cells were seeded in a 10 cm dish, and 24 hours later they were transfected with 15 µg of HIV Gag/Pol-expression plasmid (pCMVΔ8/9), 7.5 µg of Env-expression plasmid (pVSV-G) and 15 µg of shRNA expressing plasmid (pLKO-GFP-shRNA). Polyethylenimine (PEI MAX 40000 from Polysciences) was used for packaging of the lentivirus vectors. 48 hours after transfection, the supernatant was collected and concentrated using ultracentrifugation. $3x10^5$ MT-1 or Jurkat cells were seeded in 12-well plates and infected with the concentrated virus.

Flow cytometry

To measure GFP expression, cells were washed once with FACS buffer (PBS+ 2%FBS) and analyzed with a FACSVerse directly or after staining with LIVE/DEAD Near-IR (Molecular Probes) for viability. For apoptosis, cells were stained with Alexa Fluor 647 Annexin V and 7-AAD (Biolegend). For intracellular staining of Tax, cells were fixed and permeabilized with the Foxp3 Intracellular Staining Buffer Set (eBioscience), and then stained with anti-Tax mouse monoclonal antibody (clone MI73) and anti-Mouse IgG Alexa Fluor 647 (Life Technologies). Cells were also stained with DAPI to exclude sub G1 dead cells.

References

1. Zahoor MA, Philip S, Zhi H, & Giam CZ (2014) NF-kappaB inhibition facilitates the establishment of cell lines that chronically produce human T-lymphotropic virus type 1 viral particles. *J Virol* 88:3496-3504.

Supplementary Information

Note S1. Stochastic simulation of intracellular Tax expression model

A previous study on HIV elegantly established a mathematical model of 5' long terminal repeat (LTR) activation by Tat in latently infected cells[1]. We adapted this model to simulate the regulation of the HTLV-1 5'LTR by Tax (See Fig. S3*A* for a schematic representation). We tested several parameters used in the HIV study and adjusted them to fit our experimental data of transient Tax expression (Fig. S3*B-D*, Table S1, and Note S2). A simple deterministic two-state model of Tax positive feedback was developed (see the chemical reaction scheme in Fig. S3*A*)[1]. A stochastic simulation model corresponding to the deterministic two-state model was constructed and implemented by the Gillespie algorithm[2] to investigate the stochastic dynamics of Tax expression, especially for the length of Tax expression (t_{period}) and the interval between Tax expression periods (t_{interval});

$$\frac{d}{dt}[5'LTR_{OFF}] = -k_{ON}[5'LTR_{OFF}] + k_{OFF}[5'LTR_{ON}] - k_{\mathrm{bind}}[5'LTR_{ON}][Tax] + k_{\mathrm{unbind}}[5'LTR_{ON-Tax}],$$

$$\frac{d}{dt}[5'LTR_{ON}] = k_{ON}[5'LTR_{OFF}] - k_{OFF}[5'LTR_{ON}],$$

$$\frac{d}{dt}[mRNA] = k_m[5'LTR_{ON}] + k_a[5'LTR_{ON-Tax}] - \delta_m[mRNA],$$

$$\frac{d}{dt}[Tax] = k_p[mRNA] - k_{\mathrm{bind}}[5'LTR_{OFF}][Tax] + k_{\mathrm{unbind}}[5'LTR_{ON-Tax}] - \delta_p[Tax],$$

$$\frac{d}{dt}[5'LTR_{ON-Tax}] = k_{\mathrm{bind}}[5'LTR_{OFF}][Tax] - k_{\mathrm{unbind}}[5'LTR_{ON-Tax}].$$

The description of the chemical reaction scheme and parameter values used in our simulations is summarized in Table S1. The Gillespie direct method was coded in C++ (the source code is freely available upon request) to achieve fast and efficient computations. Each stochastic simulation was performed until a computational time of 100000. An arbitrarily chosen time-scale factor 0.00065 was multiplied with the computational time to imitate the

experimental time period for the live cell imaging (65 hours as a maximum period). Initial conditions for all chemical reaction agents were set to 0, except for $5'LTR_{ON}$ (the second variable in the deterministic model), which was set to 1. A sample path of Tax protein population (the fourth variable in the deterministic model) is presented in Fig. S3*B*.

Note S2. Fitting mathematical model to data

Simulated Tax-periods generated from the stochastic simulation model were fitted to the distribution of experimentally observed Tax-periods (black bars in Fig. S3*C*). Using the "entropy" package of statistical software R[3], we calculated the discretized Kullback-Leibler divergence for simulated and experimental data of Tax-periods with equal-sized 10 bins (7 hours each) to evaluate the closeness of fit between the simulation and experimental data. In fitting, all the parameters except for k_{bind} and k_{unbind} were kept fixed to the reference value (Table S1). A brute-force computation of closeness was performed on a uniform 16×16 grid of $(k_{\mathrm{bind}}, k_{\mathrm{unbind}})$, where the ranges of k_{bind} and k_{unbind} are set to [0.0005,0.04] and [0.005,0.04], respectively. The ranking of closeness (ranking 1-256) is presented in Fig. S3*D*.

Note S3. Agent-based simulation of cell population dynamics

The agent-based model (ABM) simulating the population dynamics of MT-1 cells transfected with shRNA as a Poisson (birth-death) process with a stage transition as described in Fig. S5*A* was constructed based on the individual-based Gillespie algorithm: an extension of the conventional Gillespie algorithm[2] which allows one to include non-exponentially distributed waiting-times of event occurrences as a non-homogeneous Poisson process[4]. Possible events include cell division, death, and stage transitions among three cell types. The ABM we constructed describes cell division and death events as a Poisson process with the growth rate parameter $g = 0.68$ for both Tax-negative cells expressing high levels of anti-apoptotic genes ($T_{\mathrm{off}}A_{\mathrm{high}}$cells) and Tax-negative anti-apoptotic gene expression low ($T_{\mathrm{off}}A_{\mathrm{low}}$) cells, and the growth rate parameter $g_{\mathrm{on}} = 0.54$ for Tax-expressing (T_{on}) cells. The lower value for the g_{on} parameter reflects our experimental observation that the transition to G2/M occurs only 80% as quickly in T_{on} cells (see Fig. 4). We used the death rate parameter $\delta = 0.25$ for $T_{\mathrm{off}}A_{\mathrm{low}}$ cells. The stage transition of Tax negative cells from anti-apoptotic gene expression high to low is also described by a Poisson process with rate parameters $p = 0.02$ and $p^* = 0.40$ for shNC and shTax4 cells, respectively. Note that the parameters governing the rates of other stage transitions (i.e. the transition from $\mathrm{Tax}_{\mathrm{on}}$ to $\mathrm{Tax}_{\mathrm{off}}$ and vice versa) need to be determined based on sampling from experimental (or previously simulated) Tax-period/Tax-interval data, which would be appropriately fitted with non-exponential distributions. In fact, our ABM allows us to include sampling of arbitrary waiting-times (Tax-periods and intervals) from experimental (or previously simulated) data. We summarize the parameter values in Table S2.

To simulate and reproduce the dynamics of shRNA-transduced cells growing in a mixture with untransduced cells, we need also to consider the growth rate of the untransduced cells. Untransduced MT-1 cells would show the same population dynamics as shNC-transduced

cells in our ABM. Numerical simulation results from our ABM exhibit transient behaviors until day 3 due to demographic stochasticity when the number of cells is small (i.e., around 10 cells). However, following this transient phase, a steady state phase is reached in which our numerical results are quite stable. Hence we started our ABM after day 3 to avoid possible artifacts due to demographic stochasticity in a transient phase. We fixed the initial number of MT-1 cells with and without shRNA at day 3 at 104 and 63, respectively. (The corresponding expected numbers of cells at day 0 are 12 and 11, respectively.) The initial conditions of ABM simulation were set as follows. For both shNC and shTax4 cells, less than 10% of the total cell population are T_{on}. However, we assumed that large and small fractions of the total cell population consist of $T_{\mathrm{off}}A_{\mathrm{high}}$ for shNC and shTax4 cells (78% and 2%), respectively. This is because we needed to consider the effect of Tax knockdown on the expression of anti-apoptotic genes. We also confirmed that the choice of these values is not qualitatively essential in a steady state phase (results not shown). ABM simulations for mixtures of MT-1 cells with and without shRNA were performed until day 21. Simulation results with Tax-period sampling from experimental and simulated values for shNC and shTax4 cells are shown in Fig. 5 and Fig. S6, respectively.

Since the total number of MT-1 cells without shRNA reaches 10^4 around day 10, the propensities of cell division, death and transition rarely change. In other words, the propensities required to determine the relative occurrence of an event to be selected in the Gillespie algorithm remain almost constant for a period of computational time. This allows us to skip updating the propensities, which efficiently reduces computational time. Hence we adopted an approximation algorithm to update propensities only once every 10000 events if the total cell number exceeds 10^4 (after day 10), and confirmed that this approximation does not affect population dynamics (results not shown). The ABM is coded and implemented by statistical software R. (The source code is available upon request.)

Note S4. Survival probability for subpopulations

We performed a tracking simulation for the shNC population to investigate the difference in survival time among subpopulations. Note that in this simulation, subpopulations are defined as Tax-experienced (cells in which Tax has already been expressed at least once) or Tax-naïve (cells in which Tax has never been expressed). Let a denote the age of a daughter cell, which corresponds to the time since a mother cell divided. Since we were calculating survival probabilities, we only needed to track the survival of one cohort: a subgroup of cells with the same age a. In the tracking simulation, we recorded the times of cell death and the statuses of all cells, and produced probability density distributions $f_i(a)$ for the survival time of the Tax experienced $(i = e)$ and naïve $(i = n)$ subpopulations. Survival probabilities are then derived as $F_i(a) = 1 - \int_0^a f_i(s)ds$[5]. In this tracking simulation, we used the same parameter values that we employed in the ABM simulations (Table S2 and Note S3), except for the growth rate ($g = 0$ for the tracking simulation because tracking requires only one cohort) and the initial number of a target cohort which is set to 10000. Survival probability curves of the two subpopulations are shown in Fig. 5*G*: about 90% of Tax non-experienced cells die within two weeks, but Tax-experienced cells survive much longer, suggesting that Tax expression can make a significant contribution to cell survival.

Supplementary References

1 Razooky, B. S., Pai, A., Aull, K., Rouzine, I. M. & Weinberger, L. S. A hardwired HIV latency program. *Cell* **160**, 990-1001 (2015).

2 Gillespie, D. T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics* **22**, 403-434 (1976).

3 R Core Team. R: A Language and Environment for Statistical Computing, <http://www.R-project.org/> (2015).

4 Nakaoka, S. & Aihara, K. Stochastic simulation of structured skin cell population dynamics. *J Math Biol* **66**, 807-835 (2013).

5 Aalen, O., Borgan, O. & Gjessing, H. *Survival and event history analysis: a process point of view*. (Springer Science & Business Media, 2008).