



**HAL**  
open science

## Murine leukemia virus (MLV) P50 protein induces cell transformation via transcriptional regulatory function

Charbel Akkawi, Jerome Feuillard, Felipe Leon Diaz, Khalid Belkhir, Nelly Godefroy, Jean-Marie Peloponese, Marylene Mougel, Sebastien Laine

### ► To cite this version:

Charbel Akkawi, Jerome Feuillard, Felipe Leon Diaz, Khalid Belkhir, Nelly Godefroy, et al.. Murine leukemia virus (MLV) P50 protein induces cell transformation via transcriptional regulatory function. *Retrovirology*, 2023, 20 (1), pp.16. 10.1186/s12977-023-00631-w . hal-04235300

**HAL Id: hal-04235300**

**<https://hal.science/hal-04235300>**

Submitted on 10 Oct 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

RESEARCH

Open Access



# Murine leukemia virus (MLV) P50 protein induces cell transformation *via* transcriptional regulatory function

Charbel Akkawi<sup>1</sup>, Jerome Feuillard<sup>1</sup>, Felipe Leon Diaz<sup>1</sup>, Khalid Belkhir<sup>2</sup>, Nelly Godefroy<sup>2</sup>, Jean-Marie Peloponese<sup>3</sup>, Marylene Mougel<sup>1\*</sup> and Sebastien Laine<sup>1\*</sup>

## Abstract

**Background** The murine leukemia virus (MLV) has been a powerful model of pathogenesis for the discovery of genes involved in cancer. Its splice donor (SD)-associated retroelement (SDARE) is important for infectivity and tumorigenesis, but the mechanism remains poorly characterized. Here, we show for the first time that P50 protein, which is produced from SDARE, acts as an accessory protein that transregulates transcription and induces cell transformation.

**Results** By infecting cells with MLV particles containing SDARE transcript alone (lacking genomic RNA), we show that SDARE can spread to neighbouring cells as shown by the presence of P50 in infected cells. Furthermore, a role for P50 in cell transformation was demonstrated by CCK8, TUNEL and anchorage-independent growth assays. We identified the integrase domain of P50 as being responsible for transregulation of the MLV promoter using luciferase assay and RTqPCR with P50 deleted mutants. Transcriptomic analysis furthermore revealed that the expression of hundreds of cellular RNAs involved in cancerogenesis were deregulated in the presence of P50, suggesting that P50 induces carcinogenic processes *via* its transcriptional regulatory function.

**Conclusion** We propose a novel SDARE-mediated mode of propagation of the P50 accessory protein in surrounding cells. Moreover, due to its transforming properties, P50 expression could lead to a cellular and tissue microenvironment that is conducive to cancer development.

**Keywords** Retrovirus, Murine leukemia virus, SD, P50, Cancerogenesis, Leukemia, Transcription regulation, Transcriptomic

\*Correspondence:

Marylene Mougel  
marylene.mougel@irim.cnrs.fr  
Sebastien Laine  
sebastien.laine@irim.cnrs.fr

<sup>1</sup>Team R2D2: Retroviral RNA Dynamics and Delivery, IRIM, UMR9004, CNRS, University of Montpellier, Montpellier, France

<sup>2</sup>ISEM, CNRS, EPHE, Université Montpellier, IRD, Montpellier, France

<sup>3</sup>Team APIR: IRIM, UMR9004, CNRS, University of Montpellier, Montpellier, France



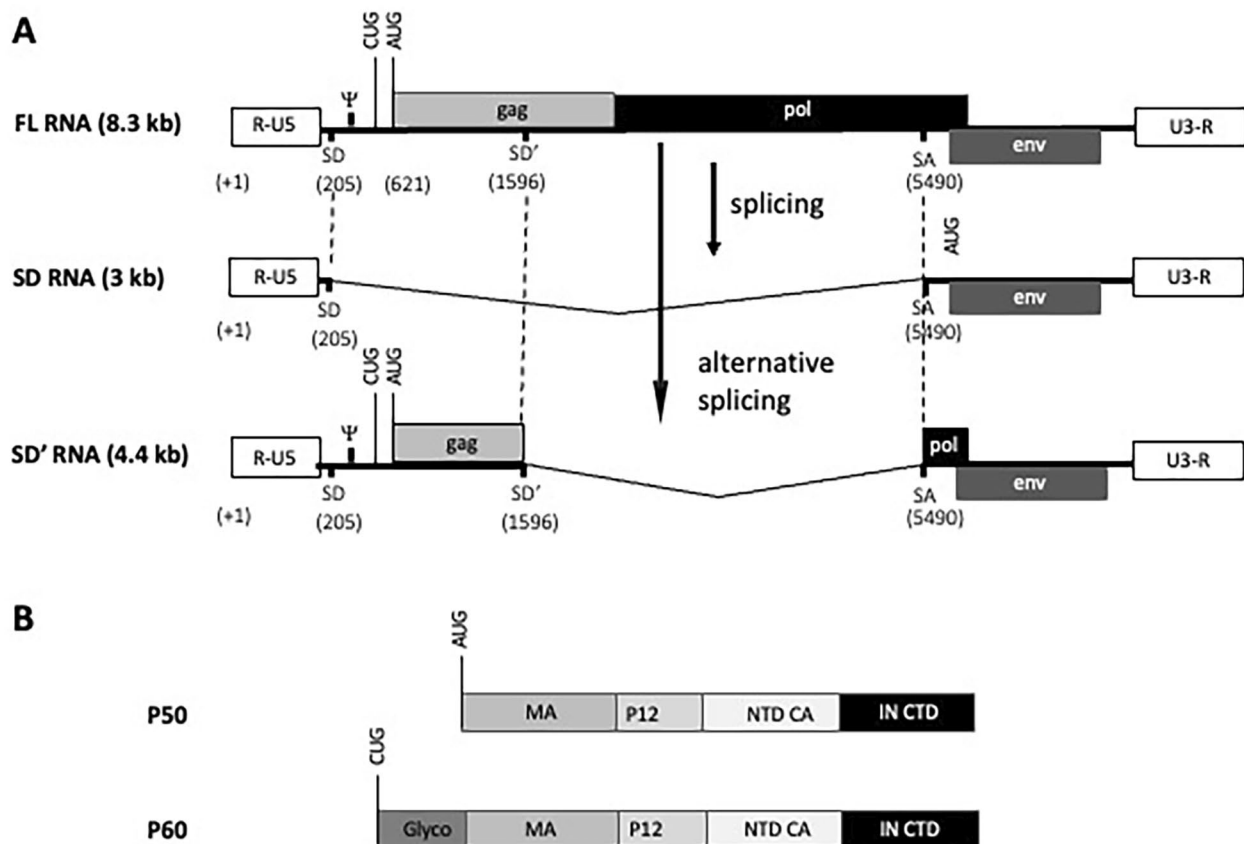
© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

### Introduction

All retroviruses share an elementary genomic organization with the *gag*, *pol* and *env* genes. They are divided into complex and simple families according to the complexity of their genomes. This complexity comes from the number of alternative splicing events of the retroviral genomic RNA, also called full-length RNA (FL RNA). In complex retroviruses, these splicing events (~30) lead to the synthesis of numerous regulatory and accessory proteins that are essential for the viral cycle, whereas simple retroviruses undergo only 1–2 splicing events. The links between retroviruses and cancer are now well documented. Both complex and simple retroviruses were found to be involved in cancer development. For complex retroviruses such as the Human T Leukemia virus 1 (HTLV-1), the viral protein Tax and HBZ proteins are crucial for cell transformation [1]. The cell transformation mechanisms are different for simple retroviruses. For instance, the Rous Sarcoma virus (RSV) includes *v-Src*, a viral-oncogene (*v-onc*) in its genome, which is essential for cancer development [2]. When simple retroviruses do not encode oncogene, the current model proposes that

tumors arise following proviral integration near a cellular proto-oncogene, a phenomenon termed proviral insertional mutagenesis [3]. Viral integration deregulates the expression of cellular genes involved in cancerogenesis. The simple virus, Murine Leukemia Virus (MLV), induces tumors by insertion of its proviral DNA in regions near cellular oncogenes, thereby deregulating their expression [4, 5]. MLV insertional mutagenesis has been particularly valuable in deciphering molecular mechanisms of hematopoietic cancers, and for the identification of new cellular proto-oncogenes [6–10].

Two strains of MLV, the simple prototypic retrovirus, are well-studied: Moloney (Mo-MLV) and Friend (F57). FL RNA of both MLVs encodes the structural Gag proteins and undergoes two splicing events (Fig. 1A): the canonically spliced SD mRNA encoding Env proteins and the alternatively spliced SD' mRNA encoding the proteins P50 and P60. The subgenomic 4.4 kb SD' mRNA is derived from a splice donor site (called SD') within the MLV *gag* sequence and the canonical splice acceptor site (SA) located in *pol* gene [11] (Fig. 1A). SD' was shown to be important for MLV biogenesis as an inactivating



**Fig. 1** MLV RNAs and proteins encoded by SD' RNA. **(A)** RNAs produced by MLV. Positions are numbered from cap site (+1) of the MLV gRNA. Non-coding elements R-U5 and U3-R are given at 5' and 3' ends, respectively, as well as *gag*, *pol* and *env* genes. Ψ depicts the packaging signal. Donor and acceptor splicing sites, SD, SD' and SA, are indicated. The translation initiation sites of P60 and P50, CUG and AUG, respectively, are shown. **(B)** SD' RNA translation. The different domains composing P50 and P60 are boxed. The N-terminal of CA and the C-terminal of IN are called NTD CA and CTD IN, respectively

mutation of SD' (without change in the *gag* amino acid sequence) reduces viral replication in murine cells [11]. We showed previously that SD' RNA shares similar abilities with FL RNA throughout the viral life cycle. SD' can be detected as a cDNA copy integrated into the host genome, demonstrating that that SD' RNA is reverse transcribed and consequently behaves as a retroelement called SDARE for spliced donor-associated retroelement [12]. Moreover, SD' RNA is efficiently packaged into progeny virions. A detailed analysis of the RNA content of MLV particles showed that SD' RNA can form SD'/FL heterodimers [13, 14]. Since SD'/SD' has been detected in vitro [13], we hypothesized that virions might contain only SD'/SD' homodimers, without FL RNA. Thus, we tested whether particles containing SD' RNA only could enter the cell and produce P50 protein. We have studied the role of P50 in this context as well as in MLV-infected cells.

Studies in vivo revealed that inoculation of a Mo-MLV mutant (MSD1/M1) harboring an inactive SD' site in newborn Swiss mice changes the MLV leukemogenic properties [15]. Indeed, whereas WT MLV exclusively induces T lymphoma, the MSD1/M1 mutant can induce a broad panel of leukemia (T lymphoma, erythroleukemia, myelomonocytic leukemia and other leukemia with undetermined origin). The same mutant was analyzed in the model of MLV-induced myeloid leukemia (MML) in pristane-treated BALB/c mice [16]. In this model, SD' is used for the production of an oncogenic gag-myb fusion RNA. However, we do not know whether the implication of SD' splicing site in tumorigenesis might also involve SD'-encoded proteins (Fig. 1B). Indeed, SD' mRNA encodes two proteins, P50 and P60, initiating at two distinct initiation codons in the same reading frame: AUG<sup>gag</sup> and CUG<sup>glyco-gag</sup>, also used for the synthesis of Gag and Glyco-Gag polyproteins, respectively [17–20]. Like Glyco-Gag, P60 is expressed at lower level than P50 [12]. The P50 protein harbors the N-terminal Gag domain including the Matrix (MA), P12 and the first 110 amino acids of Capsid (CA) domain in frame with the last 116 amino acids of integrase (IN) domain from *pol* gene (Fig. 1B). P60 possesses the same domains with 88 extra amino acids in the N-terminal region. We showed that P50 fused to GFP localizes around the nucleus and mainly at the plasma membrane and is packaged in virions suggesting a role during early and late events of the MLV life-cycle [12]. In 2020, S. Ross's team reported for the first time the role of P50 in counteracting the cellular restriction factor APOBEC3 in mice infected by MLV [21]. Like the HIV-1 auxiliary protein Vif, P50 overcomes APOBEC3 restriction by preventing its packaging into virions but without inducing mouse APOBEC3 degradation. APOBEC3 resistance was also ensured by Glycogag

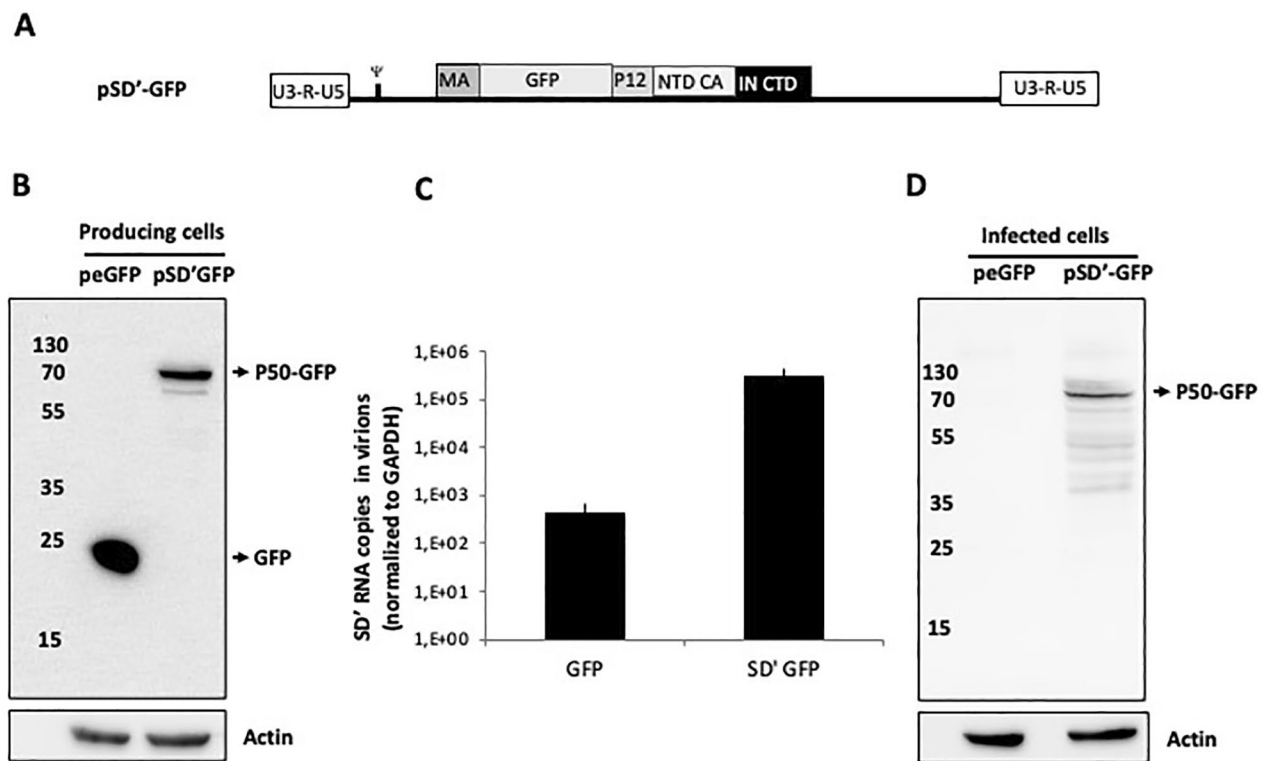
by limiting its access to the reverse transcription complex [22–24].

The perinuclear localization of P50 remains intriguing. P50 harbors the P12 and C-terminal domain (CTD) of IN that show chromatin association capacity [25–28]. P12 contains a nucleosomal-binding domain required for the association of MLV pre-integration complex to the host cell DNA during infection [27, 28]. The IN-CTD domain contains a BET interacting domain, which guides integration targeting by tethering the pre-integration complex near transcription start sites [29, 30]. All these elements suggest that P50 could also play a role in the nucleus. To this end, we have undertaken investigations into the putative role of P50 in carcinogenic processes and have discovered a transcriptional regulatory function of P50. Finally, we have performed a transcriptomic analysis to identify cellular genes whose expression is altered by the presence of P50.

## Results

### P50 can spread from cell to cell via viral particles harboring the SD' retroelement

SD' RNA is selectively packaged in MLV virions. After virus entry, SD' RNA is reverse transcribed and integrated into the host genome [12, 13]. However, SD' expression, i.e. P50 production, in these infected cells has not yet been assessed. To do so, we co-transfected murine NIH 3T3 cells with plasmids expressing the SD' retroelement or GFP vector as a control RNA that is not specifically encapsidated in virions, as well as a Mo-MLV molecular clone lacking its packaging signal (pMov9.1 Psi-) to produce virions containing only SD' RNA as a genome [31]. Three days post-transfection, viral protein expression was analyzed by western blot (WB) and harvested culture media were used to infect NIH 3T3 cells. As GFP reporter gene was inserted in frame between the MA and the P12 domains in the P50 sequence (Fig. 2A), P50-GFP was detected as a 75 kDa protein by WB by using an anti-GFP antibody (Fig. 2B). Virions were purified from the supernatant and their content was analyzed by RT-qPCR showing the presence of encapsidated SD' RNA (Fig. 2C). As previously reported [31], FL Psi- RNA was undetectable with values  $\leq$  mock ( $4 \times 10^1$  cps). Then, we infected NIH 3T3 cells with purified virions. After three days of infection, we detected the presence of P50-GFP (P75) in infected cells by WB (Fig. 2D, lane SD'GFP). P75 expression was not due to an unspecific packaging into MLV virion, since GFP was not found in infected cells (Fig. 2D, lane GFP). These results showed that particles harboring only SD' RNA are infectious and the SD' retroelement is transcribed and translated into the P50 protein.



**Fig. 2** Infection with retrotransposon SD'. **(A)** Schematic representation of chimeric pSD'-GFP RNA. NIH 3T3 was co-transfected with a construct expressing GFP (peGFP) or P50-GFP (pSD'-GFP) and a molecular clone of MLV harboring a deletion of Psi (pMov Psi-). **(B)** Three days post transfection, cell extracts were analyzed by WB using anti-GFP and anti-actin antibodies. **(C)** Supernatants were collected and virus content was analyzed by RT-qPCR before being used to infect NIH 3T3 cells. **(D)** Three days post-infection cell extracts of infected cells were analyzed by WB using anti-GFP and anti-actin antibodies

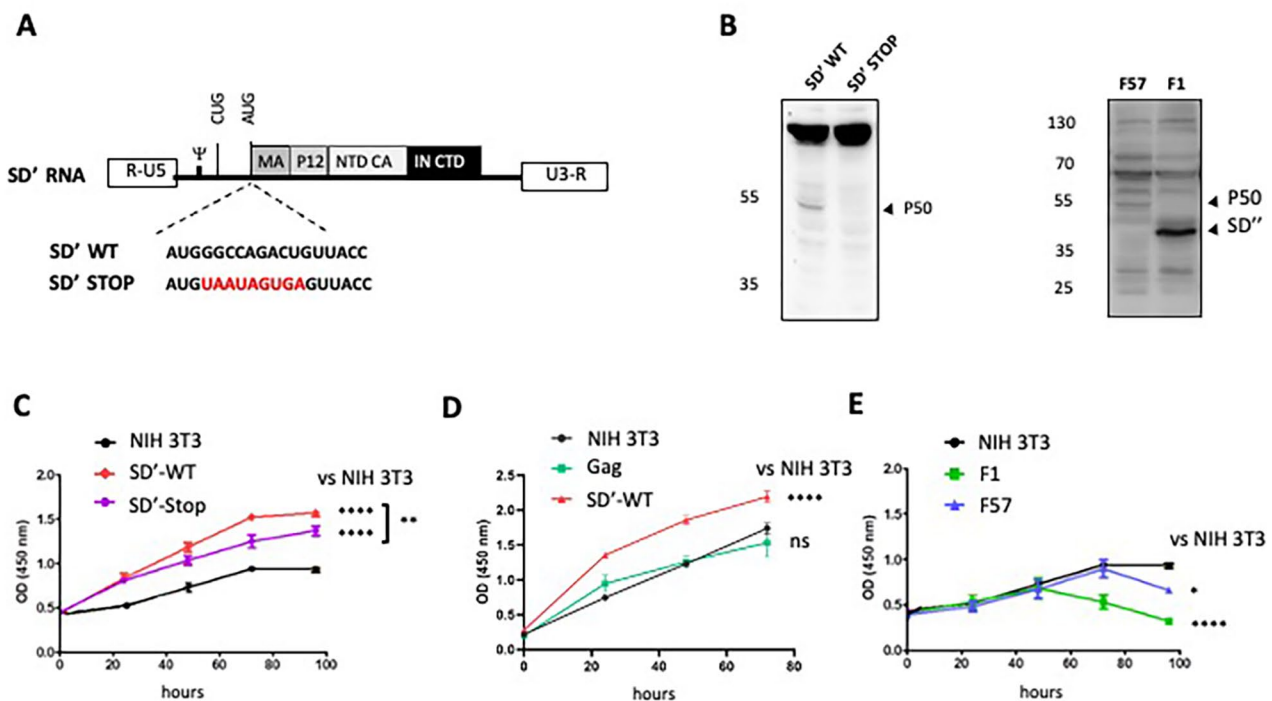
### P50 induces cell proliferation and transformation

Having shown that P50 alone can be expressed, it is interesting to investigate its impact on the host cells. We previously showed that SD' splice site is required for myeloid tumor development in vivo [16]. Nevertheless, nothing is known concerning the potential oncogenicity of P50. One of the hallmarks of oncogenes is their ability to enhance cell growth. To know if P50 is involved in cancer cell progression, we firstly asked whether expression of P50 alters proliferation of immortalized cells.

We stably transfected NIH 3T3 cells with a vector expressing a mutant SD' RNA that no longer encodes P50 (Fig. 3A and B) [12]. In parallel, we also analyzed the role of P50 in infected cells. Therefore, we used NIH 3T3 cells chronically infected with either F57 or F1 mutant (not expressing P50) [11]. The absence or the presence of P50 was monitored by WB (Fig. 3B). As previously reported when SD' splicing is inactive, an SD' protein is artificially produced with a larger NTD-CA that is no longer in phase with the IN sequence [11]. To monitor proliferation, cells were plated in a low-serum medium and analyzed daily for five days. There was a modest but significant effect of the SD' RNA (Fig. 3C) and a very small additional increase of cell proliferation due to P50 protein.

As the impact of SD' RNA could be due to the presence of SD' site or the influence of LTR promoter after insertion in cellular genome, we tested the effect on cell proliferation of the Gag gene under the control of LTR promoters since it harbors the SD' site and part of the P50 sequence except for the IN domain. This construct did not affect cell proliferation (Fig. 3D) thus confirming the importance of the full-length P50 protein in cell proliferation.

In the context of entire viruses, our results showed that F1 infected cells were less proliferative than those infected with the F57 virus (Fig. 3E). These results, compared to the proliferative effect observed when only P50 is expressed (Fig. 3C), showed that the absence of P50 inhibited the cell proliferation (Fig. 3E), although a role for SD' in inhibiting cell proliferation cannot be completely excluded. The decrease in proliferation observed in F57 infected cells was probably due to a cytopathic effect of the infection that occurs after 4 days post-infection. Taken together, these results showed that P50 increases cellular proliferation in cells. In order to know the role of P50 in MLV infection, all future experiments will be performed in the presence of P50 alone or in the viral context.



**Fig. 3** Effect of P50 on cell proliferation. **(A)** Schematic representation of SD' RNA WT or with STOP mutation (in red) inhibiting the P50 expression. NIH 3T3 cells were stably transfected with either SD' WT, SD' STOP or Gag 2LTR or chronically infected with F57 or F1 viruses. **(B)** The expression of P50 was monitored by WB, using the anti-All-Friend antibody, in F57 and F1 infected cells and in SD' WT- and SD' STOP- stably expressing cells. SD' corresponds to the product (42KDa) of a cryptic splicing that is artificially woken up when SD' alternative splicing was inactive [11]. Effect of expression of P50 or SD' RNA **(C)**, MLV Gag **(D)** or F57 or F1 viruses **(E)** on cellular proliferation was monitored daily using CCK8-coloration assay (Dojindo) (CFM-cell free medium). Experiments were performed in triplicate and the significance of differences was assessed using a one-way ANOVA test (\*\*\*\* $p \leq 0.0001$  and \* $p \leq 0.1$ , ns, not significant)

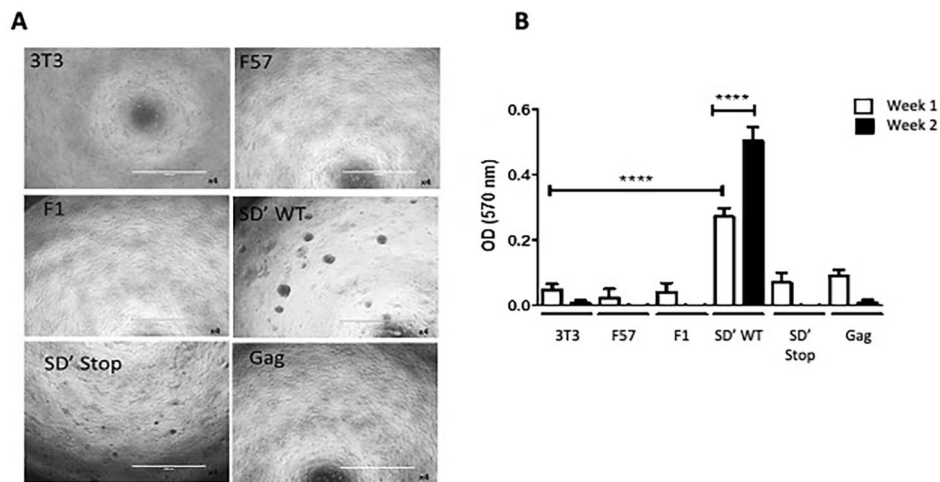
To assess the possible impact of P50 on anchorage-independent growth, which defines transforming activity in immortalized fibroblasts [32], we monitored foci formation over a two-week period of NIH 3T3 cells infected with F57 or F1 and NIH 3T3 cells that stably expressed SD'-WT or SD'-STOP mRNA, by using a 96-well plate soft agar assay (Fig. 4). No foci appear in cells carrying empty vector, Gag-plasmid, or infected by F57 or F1. We observed foci with cells expressing SD'-WT (Fig. 4A and B). However, some little foci were observed with cells expressing SD' STOP RNA but they failed to form living colonies (Fig. 4A and B), demonstrating that, contrary to P50, SD' RNA induced cell proliferation but not cell transformation (Fig. 3C). Moreover, the lack of foci observed in F57, where P50 is expressed at a low level, suggests that the effect of P50 on anchorage-independent growth is dose-dependent.

In parallel, to further investigate the mechanisms of cell transformation by P50, apoptotic activity was examined in NIH 3T3 stably expressing P50-GFP using TUNEL assay. The apoptotic level was measured by FACS analysis with cells treated or not by an apoptotic inducer. Cell expressing GFP were used as negative control. Briefly, cells were treated or not with etoposide (25  $\mu$ M) for 24 h

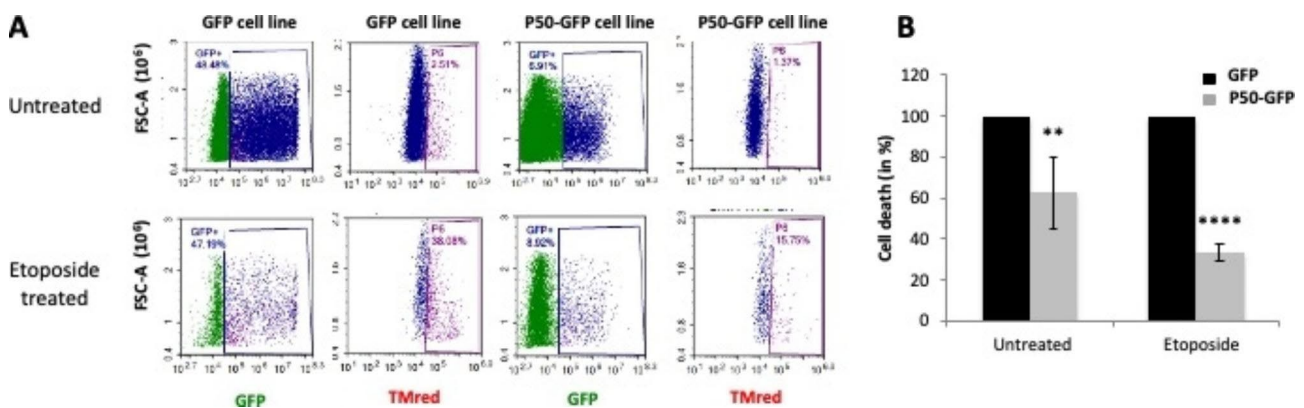
to induce apoptosis and the population of apoptotic cells was counted *via* the presence of Tm Red fluorescence revealing DNA breaking following apoptosis (Fig. 5A). Figure 5A showed that in the absence of apoptosis inducer, cells expressing P50-GFP markedly decreased the level of apoptosis (40  $\pm$  19%) compared to the GFP control. This effect is more pronounced in the presence of etoposide. Indeed, in this condition we observed a drastic decrease in cell mortality (65  $\pm$  5%). Our data (Fig. 5A and B) suggest that P50 protects the cells from apoptosis.

#### P50 is partially localized in the nucleus

P50 implication in cell transformation could suggest a nuclear function of P50. Previous fluorescence microscopy analysis reported that P50-GFP is perinuclear and accumulates at the plasma membrane [12]. To further investigate P50 subcellular localization, we analyzed by using confocal microscopy the P50-GFP location in cells transiently transfected with pSD'-GFP under the control of MLV LTR promoter, treated or not with 40 nM of Leptomycin B (LMB), a CRM1 pathway inhibitor, as described in [33] (Fig. 6A). Without treatment, P50-GFP was mainly distributed in the cytoplasm and at the plasma membrane. Interestingly, a few dots of P50-GFP



**Fig. 4** Effects of P50 on anchorage-independent growth. NIH 3T3 cells were stably transfected with SD'-WT, SD'-STOP or Gag or chronically infected with F57 or F1. Effect of P50 on anchorage-independent growth was monitored using soft agar assay. **(A)** Micrographs showing the colonies formed in the agar matrix at day 14. The scale bar corresponds to 1 mm **(B)** Absorbance-based quantitative analysis of the different cell lines. Experiments were performed in triplicates and error bars indicate SEM. The significance of differences was assessed using a one-way ANOVA test (\*\*\*\* $p \leq 0.0001$  and \* $p \leq 0.1$ , ns, not significant)



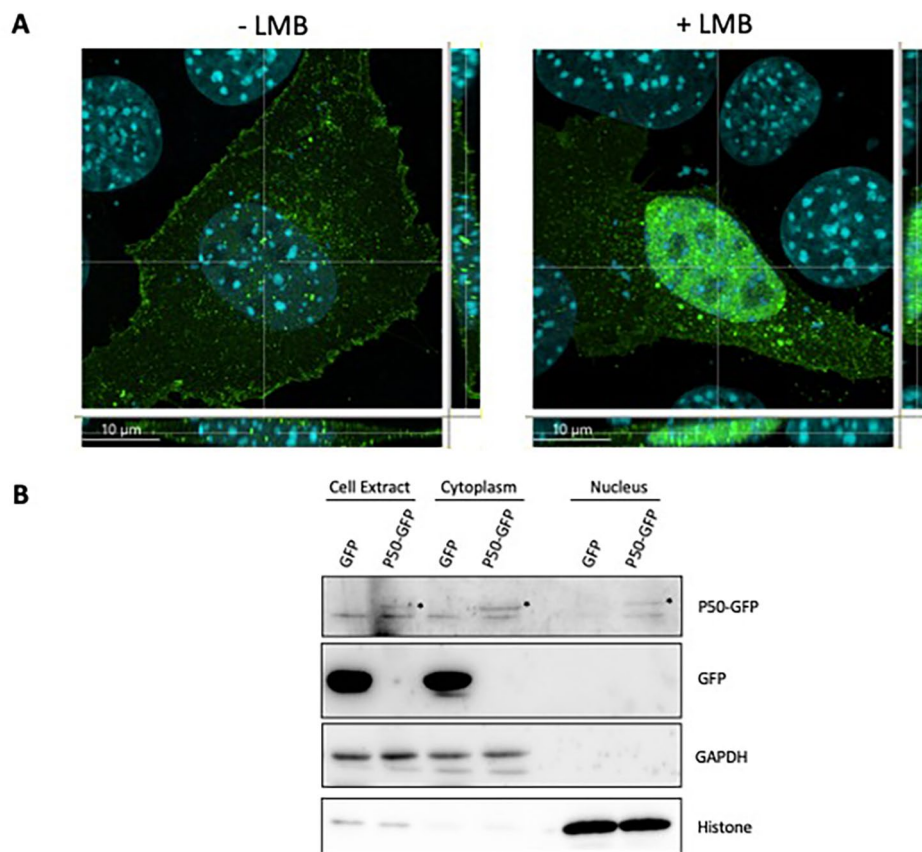
**Fig. 5** Effects of P50 on apoptosis mechanism. NIH 3T3 were transiently transfected with P50-GFP or GFP expressing plasmids and then treated or not with 25  $\mu$ M of Etoposide. **(A)** After TUNEL treatment, cells were sorted by FACS. The areas circled in blue and violet represent GFP-positive cells and among them, cells undergoing apoptosis (TMred positive), respectively. A representative experiment is shown here. **(B)** The cellular mortality was analyzed in GFP or P50-GFP expressing cells. Cells undergoing apoptosis expressing GFP are used as reference (100%). The significance of differences was assessed using an unpaired Student's *t*-test (\*\*\*\* $p \leq 0.0001$ )

were also detected in the nucleus (Fig. 6A). Treatment with LMB significantly increases the amount of P50-GFP in the nucleus (Fig. 6A). Interestingly, this increase is characterized by a greater number of dots but also by greater overall fluorescence in the nucleus. This nuclear localization was further confirmed with cell fractionation experiments (Fig. 6B). The fractionation was controlled by the detection of the cytosolic protein GAPDH and the nuclear histone H3 protein. These controls confirmed that the weak signal of P50-GFP in the nuclear fraction was not due to a leakage from the cytoplasmic fraction.

**P50 is a transregulator of the MLV transcription**

The presence of P50 in the nucleus as well as the presence of a BET binding domain in CTD-IN domain and

a nucleosomal-binding domain in P12 domain (Fig. 1B) suggest that P50 could regulate MLV transcription. Thus, we initially studied the impact of P50 on the transcriptional activity of the MLV LTR. Like HTLV-1, MLV harbors bidirectional activity of its promoter meaning that MLV RNA is transcribed from the 5'LTR and also to a lesser extent from the 3' LTR [34, 35]. In order to monitor the effect of P50 on sense and antisense transcription, we used the reporter vector MLV-Luc bidir, expressing the Renilla luciferase under the control of MLV LTR promoter in the sense orientation and the Firefly luciferase in the antisense direction [34]. NIH 3T3 cells were cotransfected with MLV-Luc bidir, and increasing concentrations of p57cDNASD' vector, expressing untagged P50. Two days after transfection, cells were collected and



**Fig. 6** Subcellular localization of P50. **(A)** NIH 3T3 cells were transiently transfected with p50-GFP (see Methods). Two-days post-transfection, cells were treated or not with 40 nM of LMB for 4 h and subcellular localization analysis was performed with confocal microscopy. Nuclei were stained with DAPI (blue) and P50-GFP appeared in green (GFP channel). Scale bar is 10  $\mu$ m. **(B)** Total cell extracts, cytoplasmic and nuclear fractions of cells transfected with p50-GFP or empty pEGFP plasmid, were analyzed by WB using antibodies against GFP, GAPDH and Histone 3. The stars indicate the position of P50-GFP.

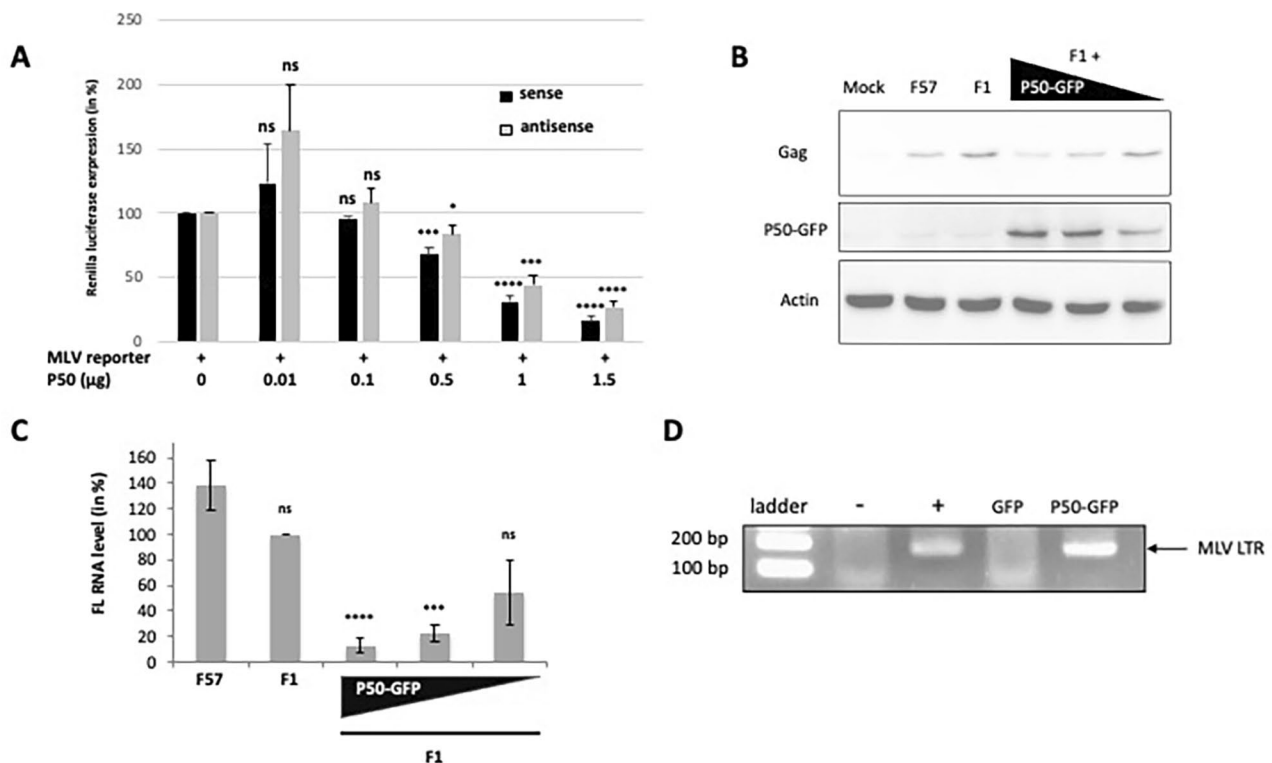
luciferase activity was measured. As shown in Fig. 7A, the lowest amounts of P50 did not significantly change the activity of MLV promoter in the sense direction (0.01  $\mu$ g: 123  $\pm$  31% and 0.1  $\mu$ g: 96  $\pm$  1% of the control level without p50cDNASD') while higher amounts (1 and 1.5  $\mu$ g) significantly reduced luciferase activity (69  $\pm$  5% and 84  $\pm$  4% reduction, respectively). While the effect on the antisense transcription was less marked, it was nevertheless significant (1.5  $\mu$ g: 73  $\pm$  5% reduction).

To study the P50 effect on MLV expression, NIH 3T3 cells were co-transfected with a F1-MLV and a vector expressing P50-GFP in dose-dependent manner. Total proteins were extracted and separated by SDS-PAGE to monitor the expression of Gag translation by WB (Fig. 7B). Interestingly, results showed a dose-dependent decrease of Gag translation along with an increase of P50-GFP level, which corroborates the results obtained with the luciferase assays (Fig. 7A). Furthermore, we monitored the levels of FL RNA by RT-qPCR (Fig. 7C). This quantification revealed a strong decrease of FL RNA (87  $\pm$  6%) in the presence of high level of P50-GFP and

a decrease of 45  $\pm$  11% was still observed at a low level of P50 (Fig. 7C). All together these results strongly suggested that P50 inhibits the transcription of the MLV promoter.

To further analyze the P50 ability to interact with MLV LTR *in cellulo*, chromatin immunoprecipitation (ChIP) assays were performed in NIH 3T3 cells co-expressing F1-MLV and P50-GFP or GFP alone as negative control. Briefly, 48 h post transfection, formaldehyde crosslinking was performed and DNA fragments interacting with P50-GFP or GFP proteins were isolated with magnetic beads coated with GFP antibodies. Bound DNA was identified by PCR using specific probes targeting LTR promoter (Fig. 7D). A band corresponding to the expected size on LTR promoter was obtained with P50-GFP (lane 4), and not with the GFP control (lane 3). This result suggests that P50 allows co-immunoprecipitation of the LTR promoter. However, the specificity for the LTR is not exclusive since the P12 domain can interact with nucleosomes and therefore the entire MLV genome might be also pulled down.



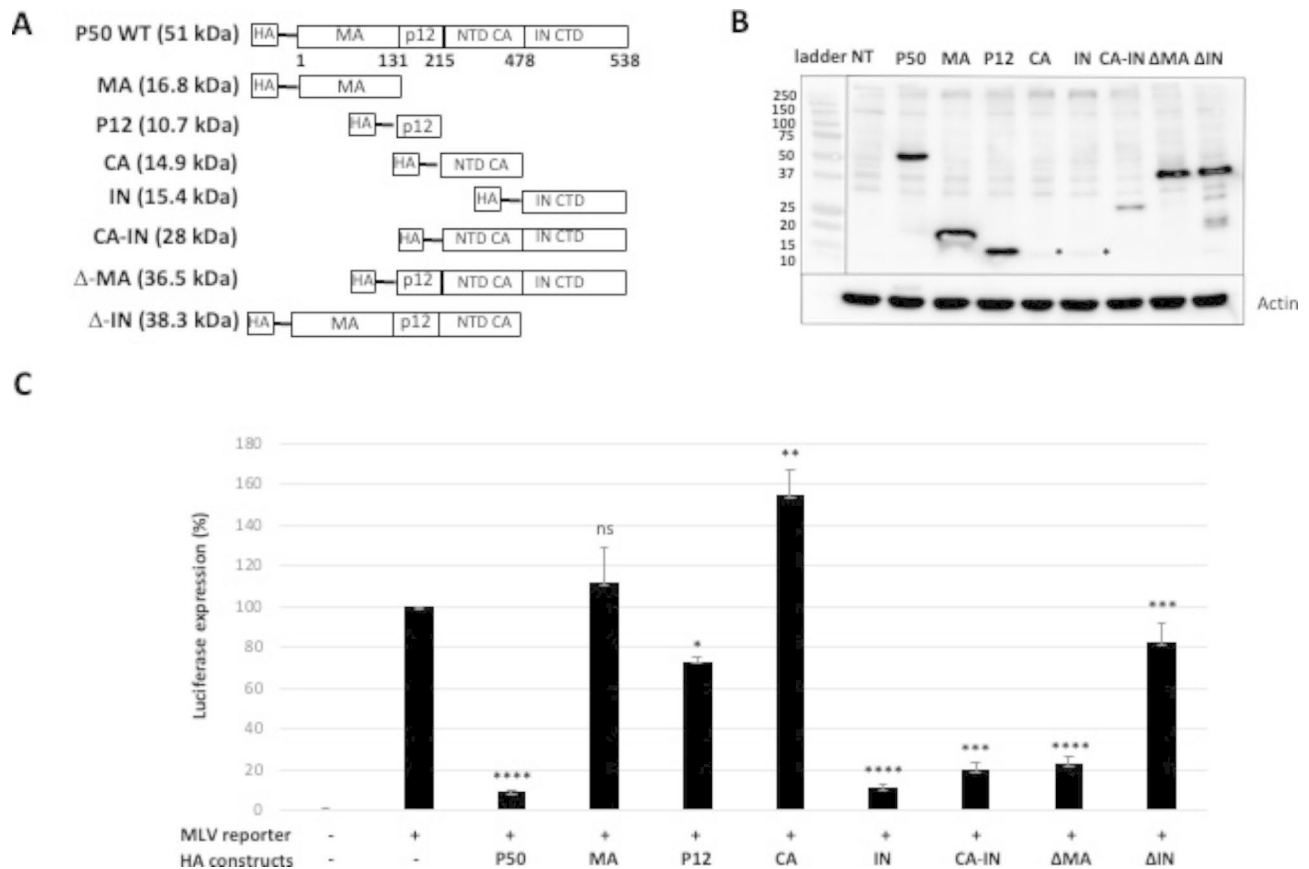


**Fig. 7** Trans-regulator activity of P50 on MLV promoter expression. **(A)** MLV promoter expression was monitored with luciferase assays. The NIH 3T3 cells were co-transfected with 1 µg of bidir MLV-LTR plasmid and increasing amounts of p57cDNASD' construct (1.5 to 0.01 µg) expressing untagged P50. Luciferase activity was analyzed 48 h post-transfection and normalized to protein concentration, measured by the Bradford assay using BSA as a standard. The activity observed with the MLV reporter alone was used as a reference (100%). Experiments were performed in triplicates. Average luciferase activities (arbitrary units) for sense and antisense were 171,639 and 4122, respectively. Values were compared using unpaired Student's t-test (ns: not significant,  $p \leq 0.1$  and \*\*\*\* $p \leq 0.0001$ ). **(B)** NIH 3T3 cells were cotransfected with plasmids expressing the mutant F1-MLV and p50-GFP and 4 8 h post transfection, Gag or P50-GFP expressions were monitored in cells and supernatants by WB analysis and **(C)** cellular MLV FL RNA was quantitated by RT-qPCR. **(D)** Chromatin immunoprecipitation of P50. NIH 3T3 cells were co-transfected with plasmids expressing F1 and P50-GFP or GFP alone. Two days post transfection cell extracts were collected and used for a ChIP experiment with an anti-GFP antibody. DNA fragments co-precipitated with P50-GFP were analyzed by PCR using probes recognizing MLV promoter. The plasmid encoded MLV was used as PCR positive control (+) and cell extract from untransfected cells was used as PCR negative control (-)

### IN-CTD domain contributes to transcription regulation ability of P50

P50 shares domains with Gag or Pol proteins (Fig. 1). Interestingly, in addition to the MA and P12 domains, P50 harbors the chimeric CA-IN domain formed by an in-frame fusion of the N-terminus of CA with the C-terminus of IN. In the presence of MLV protease, the complete maturation of P50 results in the chimeric CA-IN protein. We performed luciferase assays to identify which domain of P50 contributes to the regulation of LTR transcription. Multiple combinations of domains were constructed and tagged with HA in N-term (Fig. 8A). Expression of these truncated constructs were monitored by WB analysis. All mutants were well expressed except the CA and the IN domains which were weakly expressed (Fig. 8B). These constructs were co-transfected with the MLV-luc reporter in NIH 3T3 cells and Renilla luciferase activity was quantified (Fig. 8C). The results showed that P12 domain slightly decreased the luciferase activity (27

+/- 2%) whereas MA domain had no effect. Unexpectedly, the CA domain increased the luciferase activity (55 +/- 12%). Interestingly, the IN domain despite its low expression level, showed similar transrepression activity (89 +/- 2%) to that observed with entire P50 protein (Fig. 8C). Thus, IN domain acts as the functional domain of P50 that represses MLV LTR expression. Conversely, deleted mutants were constructed to test the role of IN in transcriptional repression. The results confirmed the role of IN in MLV LTR regulation. Indeed, the  $\Delta$ IN lead to the highest luciferase activity (87 +/- 6%) among the deletion mutants. However, the  $\Delta$ IN did not reach the transcription level of the reporter alone suggesting that another domain could also contribute to the transregulation of P50. The P12 domain is probably the best candidate as it slightly decreased luciferase activity (Fig. 8C).



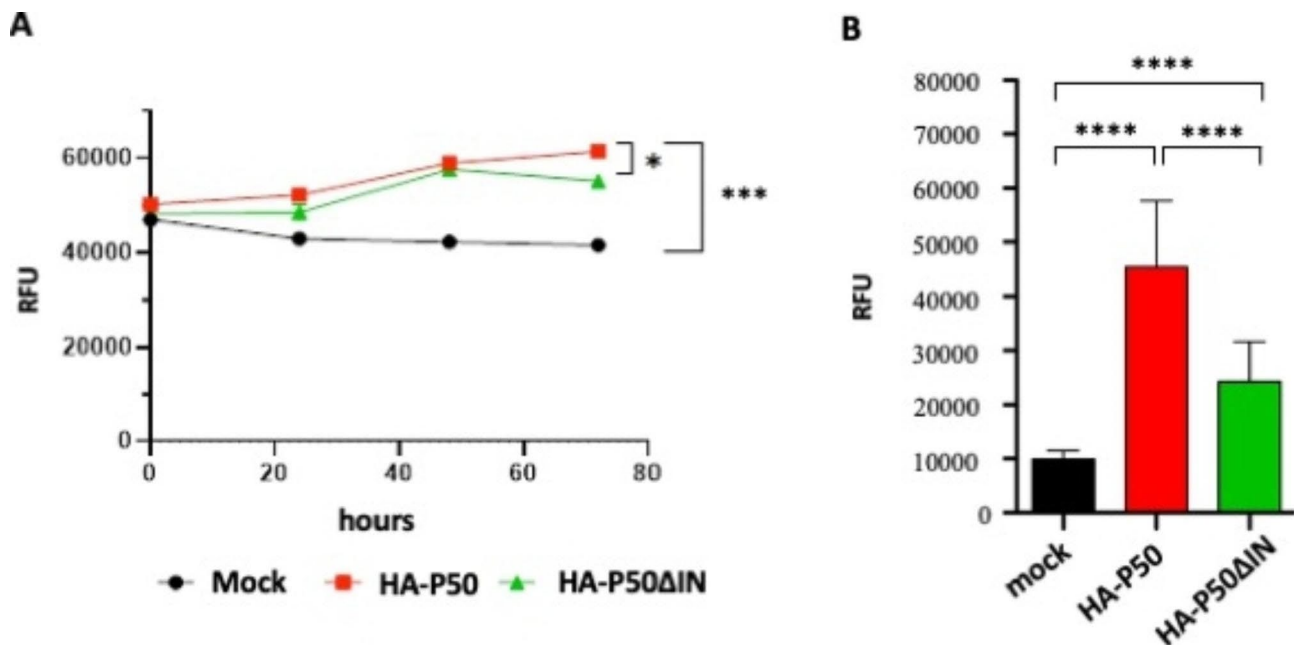
**Fig. 8** Effects of P50 domains on MLV promoter activity. **(A)** Schematic of the different HA-P50 constructs. The amino acid positions of P50 domains are given in bold. The molecular weights are in brackets. **(B)** Expression of the different HA-P50 constructs was checked by WB analysis with anti-HA antibody. The stars show the localization of N-term CA and C-term IN domain. Experiments were performed in triplicates. **(C)** MLV promoter expression. NIH 3T3 co-transfected with bidir MLV-LTR and pRK5-HA-P50 or empty pRK5-HA constructs. Luciferase activity was analyzed 48 h post-transfection and values obtained with only bidir MLV-LTR plasmid was referred as 100%. Experiments were performed in triplicates and the significance of differences was assessed using an unpaired Student's *t*-test (ns: non significant, \*\*\*\* $P \leq 0.0001$ , \*\*\* $P \leq 0.001$ , \*\* $P \leq 0.01$  and \* $P \leq 0.1$ )

### IN domain of P50 is involved in the transforming capacity of P50

To further investigate the role of the IN domain in P50-induced cell transformation, we performed cell proliferation and growth experiments on soft-agar with NIH 3T3 cell lines stably expressing HA-P50ΔIN (Fig. 9). To monitor proliferation, cells were plated in a low-serum medium and analyzed daily for three days. The results showed that, unlike HA-P50, the presence of HA-P50ΔIN had less impact on cell proliferation (Fig. 9A). In parallel we also tested the impact of HA-P50-ΔIN on anchorage-independent growth. As previously described, we monitored the cell viability over a two-week period of NIH 3T3 cells stably expressing HA-P50 and HA-P50ΔIN mRNA, using a 96-well plate soft agar assay (Fig. 9B). Interestingly, we observed that HA-P50ΔIN had less impact on anchorage-independent growth. Taken together these results suggest that the effect of P50 on cell proliferation and anchorage-independent growth is dependent on its IN domain.

### P50 regulates transcription of cellular genes

Next, we wondered if the P50 transcriptional regulation activity could be extended to cellular genes. For this purpose, a comparative transcriptomic study was performed with NIH 3T3 (mock), NIH 3T3 chronically infected with F57 and NIH 3T3 stably expressing P50 (Fig. 10). A total of 83,173 transcripts showed a measured level of expression under the different conditions. The comparison between MLV infected cells and P50 stably expressing cells to identify differentially expressed genes, using DESeq2 under iDEP with  $FDR < 0.001$  and a fold change  $> 8$ , identified 165 upregulated and 185 downregulated transcripts (list on request). Figure 10 shows the expression profile of the regulated genes in all conditions. In infected cells (MLV), we found 92 up-regulated RNA and 72 down-regulated RNAs. The same proportion of regulated RNAs are found up-regulated and down-regulated, 99 and 88 respectively, in P50 stably expressing cells (P50) [see additional file 1 and 2 respectively]. Among these genes, 29 gene expression profiles observed



**Fig. 9** Effect of IN domain on cell proliferation and anchorage-independent growth. NIH 3T3 transfected by pRK5-HA-P50 or pRK5-HA-P50ΔIN and then stably expressing HA-P50 or HA-P50ΔIN were performed. **(A)** Cell proliferation was monitored daily using Presto Blue-coloration assay (Invitrogen) (CFM-cell free medium). **(B)** Effect of HA-P50 and HA-P50ΔIN on anchorage-independent growth was monitored using soft agar assay. Absorbance-based quantitative analysis of the different cell lines grown in the agar matrix at day 14 was performed with Presto Blue-coloration assay (Invitrogen). Results are given in Relative Fluorescence Units (RFU). Experiments were performed in triplicates and error bars indicate SEM. The significance of differences was assessed using a one-way ANOVA test (\*\*\*\* $p \leq 0.0001$  and \* $p \leq 0.1$ , ns, not significant)

in MLV and in P50 varied in a similar way. Interestingly, most of these genes are linked with cancer processes (Pld1, Fam13c, Svbpl, TMem192, Zc3h7a, RTP4, Naa30, Zgrf1, Slc33a1, Dnmt3b, Map6, Plin4, Eps8L2, Alg12, Capg and Tdp2) [36–52]. Moreover, we found that some genes are also important for leukemia development (Jarid2, Metap2, Jak3, Pla2G15, Zfp384, Casp8 and Dpf2) [53–59].

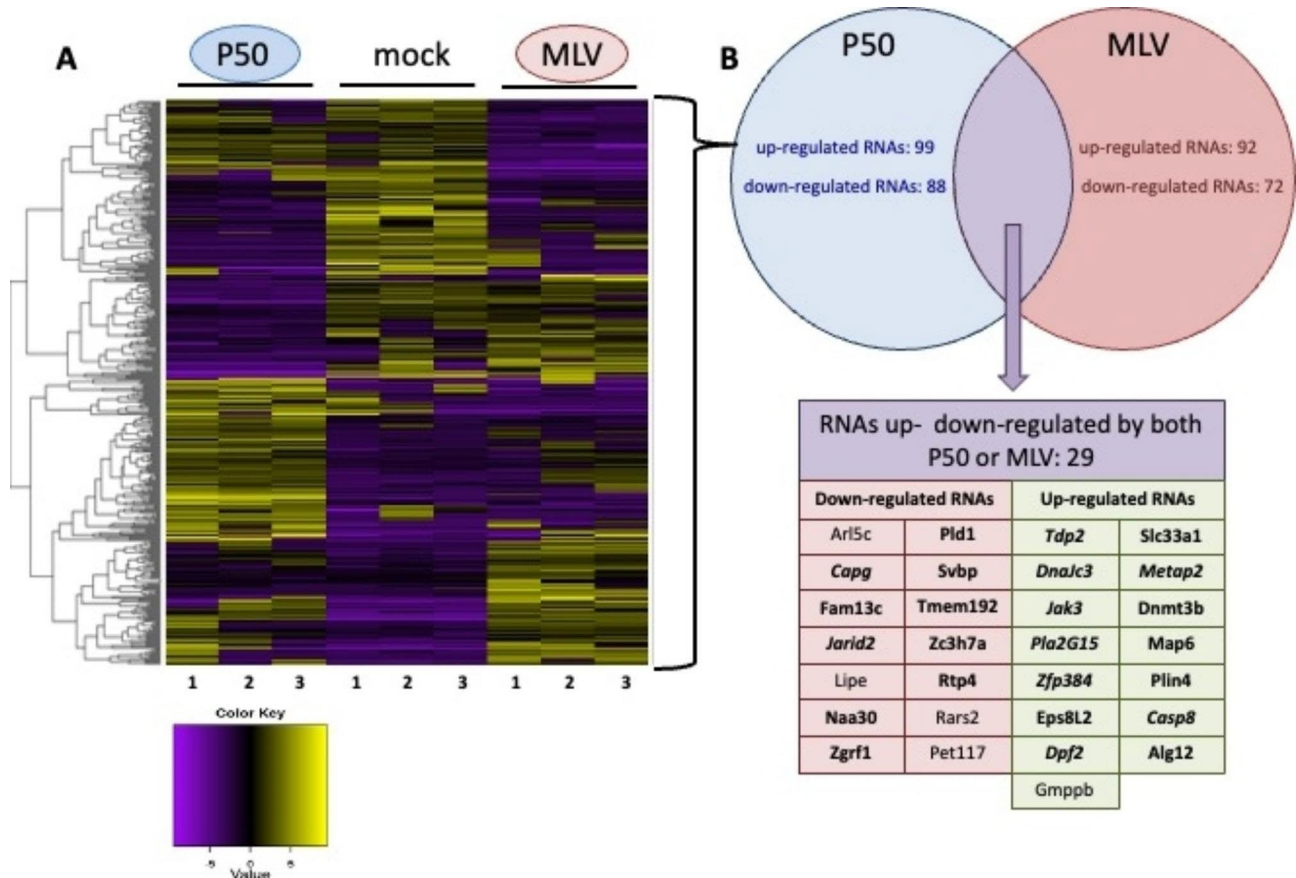
The transcriptomic analysis revealed that P50 modulated some genes involved in cancer biogenesis which is in good correlation with the effects of P50 on cell proliferation and transformation processes (Figs. 3, 4 and 5). Taken together, these results revealed that P50 is involved in cancerogenesis by acting as a transcription regulator.

## Discussion

As previously reported, the SD' RNA acts as a splice donor-associated retroelement (SDARE). Indeed, it is selectively packaged in infectious particles and after cell entry, it is reverse transcribed by the viral RT and integrated as proviral SD' by the IN [12]. Packaging is not a particularity of FL RNA and viral spliced RNAs also dimerize (for review, [60]) and are specifically packaged in virions [61]. The SD' RNA can dimerize [13] and since it carries the Psi packaging signal, it is specifically packaged with FL RNA in infectious particles. Our team reported the existence of MLV particles harboring

heterodimers FL/SD' RNA suggesting that homodimers SD'/SD' can also be packaged [14]. Here, we produced virions containing only SD' RNA (and no FL RNA) that are still infectious since P50 is expressed in infected cells. Therefore, P50 might spread to cells neighboring the primary infected cell. In this context, P50 is produced after transcription of the proviral SD' and not from SD' mRNA originated from FL RNA splicing, suggesting that SD' retroelement can spread among cells. The presence of a retroviral protein in neighboring cells has already been reported for HIV-1 secreted Tat protein [62]. This regulatory protein is able to enter cells and disturb several cellular pathways by modifying the expression of cellular genes [63]. No such protein was reported yet for simple viruses, such as MLV. Here, we revealed for the first time that MLV uses another strategy based on the release of infectious particles harboring the retrotransposon SDARE encoding the regulatory protein P50. To our knowledge, such a mechanism has never been described in the literature.

SD' RNA and/or P50 protein contribute to MLV infectivity (F57 and Mo-MLV) and pathogenesis [11]. Although a role for P50 has recently been reported in counteracting the restriction factor APOBEC3 [21], P50's role during infection remains poorly understood. Here, we showed that P50 was also involved in transcription regulation and cell transformation. The hallmarks of



**Fig. 10** Transcriptomic analysis. (A) RNA populations from mock NIH 3T3, cells chronically infected by F57 (MLV) and cell line stably expressing P50 (P50) were analyzed in transcriptomic assays (in triplicate: 1, 2 and 3). The dendrogram representing the result of the clustering of all pairwise dissimilarities between the transcript expressions is represented on the left. The (B) RNAs regulated by P50 alone (blue circle) or in viral context (red circle) are represented and RNAs identically regulated in both conditions are represented in violet and listed in the table. RNAs from genes implied in cancer are in bold and those particularly involved in leukemia in bold italic

cancer were divided in ten biological capabilities: proliferative signaling, evading growth suppressor, resisting cell death, enabling replicative immortality, inducing angiogenesis, activating invasion and metastasis deregulating cellular energetics, avoiding immune destruction, tumor-promoting inflammation and genome instability and mutation [64]. We demonstrated that P50 ticks three of these hallmarks: proliferative signaling (Fig. 3), evading growth repressor (Fig. 4) and resisting cell death (Fig. 5), indicating that P50 harbors oncogenic properties. More specifically, we have shown that the effects of P50 on certain hallmarks of cancer (proliferative signaling and evading growth repressor) are due to its IN domain (Fig. 9). Nevertheless, the molecular mechanisms of oncogenesis remain to be fully elucidated. In agreement with this ability, a low amount the P50-GFP was found to be localized in the nucleus and its transport across the nuclear envelope is CRM1 pathway dependent (Fig. 6A). Like the regulatory proteins Tat of HIV and Tax of HTLV [65, 66], which transactivate viral transcription (for review, [67, 68]), P50 also acted as a regulatory protein.

Indeed, P50 transrepressed the viral transcription and consequently decreased the FL RNA translation (Fig. 7). This regulatory function was mediated by its IN-CTD domain with a lesser contribution from its P12 domain (Fig. 8). IN domain interacts with the cellular BET proteins that guides MLV integrations near transcription start sites [30, 69]. The involvement of the BET proteins in the activity of P50 on cellular or viral promoters will need to be further investigated. Interestingly HIV-1 IN-CTD domain contributes to viral transcription regulation by Tat [70] via interactions with TAR region [71]. P12 domain, on the other hand, harbors a nucleosomal-binding domain facilitating MLV integration [27, 72]. It is possible that the P12 domain of P50 acts in synergy with IN-CTD domain to modulate the transcription regulation activity of P50.

Until now there was no evidence of a role of P50 in oncogenic processes. Our transcriptomic study revealed for the first time that in addition to acting on MLV promoter, P50 deregulated mRNAs of several cellular genes, including genes involved in essential carcinogenesis

processes (Fig. 10). Interestingly, among the RNAs regulated by MLV or P50 alone, we found the mRNAs of the genes *Ldlr*, *Tnr*, *Mapre2*, *Fnbp1*, *Hmgcr*, *Rara* and *Camk2d*, which have been previously identified as the MLV insertion sites in different cell types [7, 9, 10, 39]. Due to these transregulatory capacities, P50 could interact directly or indirectly with transcription factors or cell promoters. This phenomenon has been reported for HTLV-1 whose onco-protein Tax affects viral RNA transcription as well as cellular RNAs leading to cell transformation [73]. Among cellular genes which are transregulated by P50, the majority encode proteins involved in cell-cycle regulation, or proteins found deregulated in cancers and leukemias. Thus, we found that the expression of several mRNAs encoding proteins involved in cell proliferation (*Jak3*, *Fam13c*, *Pla2g15*, *Arl5c*, *Naa30*, *Pld1*, *Tmem192*) is affected by the expression of P50 protein alone and in the viral context [37, 38, 40, 43, 55, 56, 74]. This ability to affect cell growth is found in other viral oncoproteins. Indeed, the NS3 and NS5A proteins of the hepatitis C virus promote the proliferation of hepatic cells via the  $\beta$ -catenin pathway [75]. Interestingly, we showed P50 altered transcription of RNAs encoding several proteins such as *Fam13c* and *Jak3*, which are transcriptional activators of the  $\beta$ -catenin gene that promotes cell growth [76, 77]. Therefore, it is quite conceivable that P50 protein increases cell proliferation *via* the  $\beta$ -catenin pathway.

Similarly, several mRNAs encoding proteins involved in growth inhibitory signals (*Jarid2*, *Metap2*) were affected by P50 [53, 78]. The dysregulation of *Jarid2* gene is observed during Kaposi sarcoma-associated herpesvirus infection and leads to cell transformation [79]. *Metap2*, on the other hand, is differentially expressed in cells derived from mice developing an acute murine leukemia [54]. Thus, all these data suggest that P50 could act on several signaling pathways leading to insensitivity to growth inhibitory signals.

Regarding genes controlling cell apoptosis, we detected variation in the expression of several mRNAs controlling cell death (*Naa30*, *Tmem192*, *Casp8*) [40, 43, 58]. As a large fraction of P50 is in the cytosol, protection from apoptosis could also be due to the interaction of P50 with cellular partners. Such inhibition of apoptosis is observed for hepatitis viruses (for review, [80]) and further investigations will be necessary to test this possibility.

Our results shed light on the pathways used by P50 to regulate mechanisms involved in loss of contact inhibition, immortalization or proliferation. However, other mechanisms of carcinogenesis are also impacted by P50 (Fig. 11), since several variations in expression of mRNAs controlling cellular energy metabolism were detected: *Pet117*, which belongs to the prostaglandin G/H synthase pathway [81, 82], *Lipe* that is involved in

the regulation of lipogenesis [83], or *Rars2* involved in the biogenesis of the phosphorylating oxidative chain [84, 85]. Genes involved in other characteristics of cancers, such as inflammation, genomic instability or escape from the immune system were also impacted by P50. We thus detected an increase in mRNAs encoding *Pld* and *Svbp* proteins, responsible for angiogenesis and inflammation [86, 87], *Tdp2* involved in genomic instability, and *Zc3h7a* involved in escape from the immune system [51, 88]. The transcriptomic study also revealed altered expression of several aforementioned genes encoding proteins involved in different distinct deregulatory circuits such as *Pld*, *Tmem192*, *Naa30* and *Lipe*.

Overall, our results showed that P50 affected the expression of cellular genes involved in cellular transformation and cancerogenesis. SD' retrotransposon and the encoded P50 protein promote viral replication that could result in a cellular and tissue microenvironment conducive to the development of cancers [89, 90]. In contrast to HIV-1 Tat protein, which is secreted from infected cells and then enter into uninfected cells [91], P50 is disseminated *via* the SD' RNA contained in infectious viral particles. To our knowledge, this kind of dissemination is unique in the retroviral family.

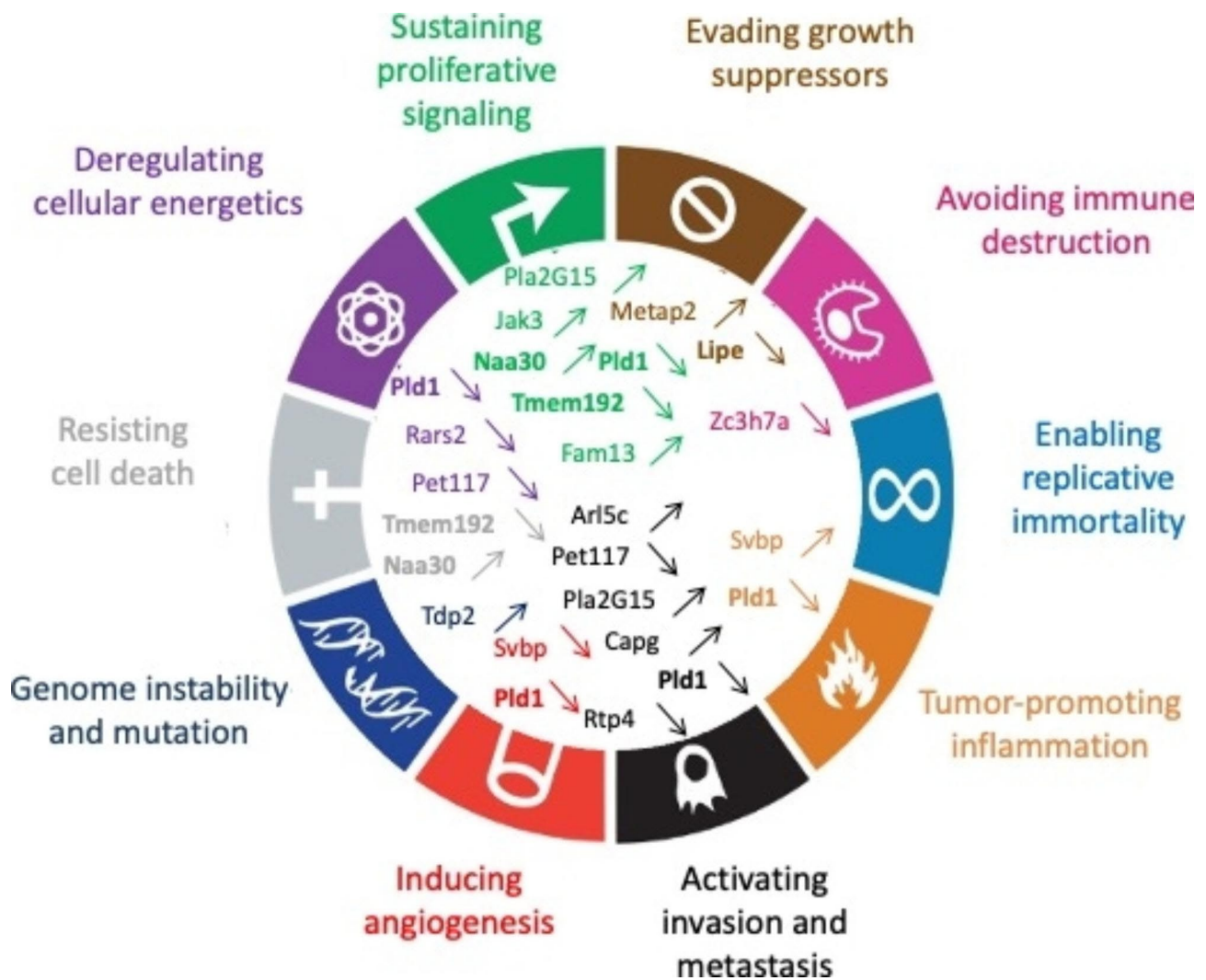
## Methods

### Plasmid constructions

Plasmids were constructed by using common cloning techniques and propagated in the *E. coli* DH5 $\alpha$  at 37°C or STBL2 strain at 30 °C for plasmids carrying LTR to prevent any recombination. Finally, all constructs were sequenced.

Viral production was based on co-expression of Mo-MLV molecular clone deleted of its Psi signal (pMov9.1 Psi-) [31] with peGFP-N2 (Clontech) as control or pBSKEco SD'GFP (pSD'GFP) vector. pSD'GFP was generated by fragment exchange using *Xho*I and *Hind*III digestion sites between pBSKEco-p12-GFP-MS2 WT [92] and pSD' [12] respectively.

For proliferation and transformation experiments, F57 strain and Friend F1 SD' mutated molecular clone (F1) are described in [11]. The p57cDNASD' vector (SD' WT), expressing SD' RNA and proteins, and the p57cDNASD' stop (SD' Stop) vector which expressing SD' RNA only were previously described [12]. The p2LTR-Gag-F57 vector (Gag 2 L), expressing Gag under LTR promoter was obtained by deleting a 4424 nucleotides fragment in pF57 molecular clone with *Sac*II and *Eco*RI enzymes. The obtained cohesive strand ends were filled with klenow enzyme and religated with T4 DNA ligase. The plasmids expressing HA-P50 (pRK5-HA-P50) and HA-P50 $\Delta$ IN (pRK5-HA-P50 $\Delta$ IN) were constructed as explained below.



**Fig. 11** P50 and cancerogenesis. Schematic of classification according to the hallmarks of cancer of different proteins encoded by altered cellular mRNAs in the presence of P50 alone and in viral context (adapted from [64]). The mRNAs encoding proteins involved in different deregulated pathways are in bold. Deregulated genes playing a role in a distinct hallmark of cancer are in same color. Upregulation and downregulation of genes are noted by upward and downward arrows, respectively

TUNEL analysis was performed with plasmid expressing P50-GFP (p1LTR-P50-GFP). p1LTR-P50-GFP was generated by fragment exchange using AflIII and HindIII digestion sites between p57cDNASD' and p57SD'/GFP described in [12]. The plasmid thus created was digested with AatII then treated with Klenow fragment and religated in order to introduce a STOP codon upstream of the ATG encoding P50.

Finally, the transactivating experiments were performed with Luciferase assays and CHIP analyses. Luciferase assays were carried out with MLV reporter vector (pAsLuc(Fire)-MLV-Luc(Reni)) containing Renilla and Firefly luciferase under LTR promoter which was generously provided by JM Mesnard [34]. Plasmids expressing P50 (p57cDNASD'), P50 tagged with HA or the multiple combinations of domains of P50 tagged with HA were

obtained by amplification with standard PCR of P50 or every domain (alone or in combination) and inserted in pRK5-HA (addgene®) using EcoRI digestion. For CHIP analyses, the Mo-MLV molecular clone (pMov) [31] was used as positive control and the plasmid p1LTR-P50-GFP expressing P50-GFP (quoted above) was used for the experiment. Details of plasmid constructions will be provided on request.

**Cell culture and transfection**

NIH 3T3, NIH 3T3 Packaging cell line [11] and NIH 3T3 chronically infected by F57 or F1 strains were grown in DMEM medium with penicillin-streptomycin (100 U/mL) and 10% foetal bovine serum (FBS) at 37°C and 5% CO2. Several stable NIH cell-lines were established by co-transfecting them with p57cDNASD', p57cDNASD'stop,

p2LTR-Gag-F57, pRK5-HA-P50 and pRK5-HA-P50 $\Delta$ IN plasmids and in a ratio of 10/1 the pGkNeo plasmid carrying the neomycin resistance. We used JetPei method according to the manufacturer's instructions (Polyplus Transfection) and selection was started two days after transfection by adding G418 drug (100 mg/mL). After 1-month, all clones were trypsinized and mixed in a new dish.

#### Western blotting

Protein analysis was performed by western blotting. Cells were scraped and lysed with the Complete lysis-M kit (Sigma) supplemented with protease inhibitor cocktail according to the manufacturer's instructions. Total protein concentration was determined by Bradford assay using a BSA standard set (Fermentas). 100  $\mu$ g of cell lysate Proteins was loaded on 12% SDS-PAGE and were electro-transferred onto nitrocellulose membrane. Gag was detected using a rat anti-capsid (CA) antibody (1/500, hybridoma H187, from B. Chesebro and M. Miyazawa). Actin was detected with an anti-actin (1/500, Sigma). P50 was detected with a Goat anti-All-Friend antibody (1/2000, from M. Sitbon). HA epitope was detected with a rabbit anti-HA antibody (1/4000, Abcam). After incubation with a peroxidase-conjugated (HRP) secondary antibody, ECL fluorescence was recorded by a CCD chemiluminescence camera system (ChemiDoc™ MP Imaging System, Bio-Rad).

#### RNA analysis by RT-qPCR

RNA was extracted from cells with TriReagent (MRC) according to the manufacturer's instructions and was quantified by RT-qPCR as described in [93]. After extraction, RNA samples were treated with RQ1 DNase (Promega) in the presence of RNasin (Promega) for 25 min at 37 °C, followed by phenol–chloroform and chloroform steps and finally precipitated with ethanol. After 70% ethanol wash, RNA pellets were dissolved in RNase free water. RNAs were quantified by measuring optical absorption at 260 nm. For RT-qPCR, the RT step was performed with Expand reverse transcriptase (Roche) and initiated with an oligo (dT) primer or the same anti-sense primer used for the qPCR reaction. 2.5  $\mu$ L of the RT reaction were engaged in qPCR reactions [93]. The following primers were used for FL or Psi+RNA: sFM76 5'-GTGGTCTCGCTGTTCCCTTGGA sense and aM275 5'-GCAGGCGCATAAAATCAGTCATAG antisense, for Psi- RNA: sMLV5356 CCCTGTACCGAGCCCG-CAACAC sense and aFM5620 GCGGACCCACACT-GTGTC antisense and for GAPDH: sGAPDH721 5'-GCTCACTGGCATGGCCTTCCGTGT sense and aGAPDH921 5'-TGGAAGAGTGGGAGTTGCTGTTGA antisense.

#### Virus production and cellular infection

A total of  $1.6 \times 10^6$  NIH 3T3 cells were co-transfected with pMov9.1 Psi- and pSD'GFP expressing P50-GFP or empty peGFP-N2 plasmid as control in 10 cm-plates with JetPei method according to the manufacturer's instructions (Polyplus Transfection). Two days post transfection, cells were collected and analyzed by Western blot, and the virus-containing media (10 ml) were collected, centrifuged 5 min at 1500xg and filtrated through a 0,45  $\mu$ M filter. The supernatant was then ultracentrifuged on a 20% sucrose/phosphate-buffered saline (PBS) cushion at 30,000 rpm for 1h30 at 4 °C. The pellet was resuspended in 400  $\mu$ L of complete DMEM. Viruses were then added to  $3 \times 10^5$  NIH 3T3 seeded the day before in a 6-wells plate in the presence of polybrene (10  $\mu$ g/mL). Two hours post infection, 2 mL of complete DMEM were added. 24 h after infection, cells were trypsinised and transfer in a new 10 cm-plate. Three days post infection, infected cells were collected and proteins were analyzed by Western blot using anti-GFP antibody and normalized with anti-actin antibody.

#### Cell fractionation

NIH 3T3 cells were seeded and transfected with p57SD' GFP or empty peGFP plasmid as described above. 48 h after transfection,  $2.6 \times 10^5$  cells were harvested and sub-cellular fractionation was conducted with PARIS™ Kit with method according to the manufacturer's instructions (Invitrogen, Carlsbad, CA, USA). In parallel,  $2.6 \times 10^5$  cells were scraped and lysed with the Complete lysis-M kit (Sigma) as described above to obtain cell extract. Half of nuclear or cytoplasmic fraction and cell extract was loaded on PAGE gel for Western blot analysis. Mouse anti-GFP (1/2000, Roche) was used as primary antibody to detect GFP and P50GFP. Cytoplasmic fractions were checked with anti-GAPDH (1/2000, Ambion) and nuclear fraction with rabbit anti-H3 (1/2000, Abcam) antibody.

#### Cell proliferation

Cell proliferation was determined using cell counting kit 8 (CCK8; Dojindo or PrestoBlue; Invitrogen) according to the manufacturer's protocol. Briefly,  $5 \times 10^4$  of each stable cell lines were plated in 96-well plates in DMEM medium supplemented with 1% heat-inactivated fetal bovine serum (FBS). 10  $\mu$ L of substract solution were added daily to each well, followed by incubation for 4 h. The cell viability in each well was determined by reading the optical density at 450 nm for CCK8 or by measuring the fluorescence intensity (RFU) with PrestoBlue at 560/590 nm on a TECAN SPARK 10 M (TECAN).

### **Anchorage-independent growth (soft agar cell culture) in 96-well plates**

Soft Agar cell culture in 96-well plates was performed according to [94]. A total of  $10^5$  stably transfected NIH 3T3 cells/well were seeded in 96-well plates with a bottom layer of 0.6% Select agar-DMEM (Invitrogen) and a top layer of 0.3% Select agar-DMEM. Fresh DMEM with 2% FCS was added to the top layer of the soft agar twice a week. The cells were allowed to grow in the humidified 37 °C incubator with 5% CO<sub>2</sub> for 1–2 weeks. Colony growth and cell viability were measured using Prestoblu (Invitrogen -A13261) by reading the OD at 570 nm or the fluorescence intensity (RFU) at 560/590 nm on a TECAN SPARK 10 M (TECAN).

### **Microscope analysis of P50-GFP**

A total of  $2.3 \times 10^5$  NIH 3T3 cells were grown on a gelatin-coated coverslip in a 6-well plate. 2 µg of pSD'GFP were transfected as described above and 48 h after transfection, cells were treated or not with 40 nM of LMB (LC Laboratories) for 4 h as described in [33]. Then, cells were fixed with Formaldehyde (3.7%) for 10 min at RT. Then, Coverslips were washed twice with PBS and then mounted with vectashield+DAPI (1/2000, Thermo Fisher Scientific). Cells were imaged on a confocal laser scanning microscope LSM980 (Zeiss) equipped with a 63x/1.4 NA oil objective and 0.2 µm z-stacking space, with a minimum of 20 stacks per cell at Montpellier Rio Imaging (MRI) facilities. Images were deconvolved with Huygens Professional (version 22.10) and, maximum intensity projections and orthogonal views were prepared in Imaris Bitplane (version 10.0).

### **TUNEL analysis**

NIH 3T3 cells were transfected with p1LTR-P50-GFP or pEGFP-N2 in a 6-well plate as described above, 24 h post-transfection,  $\pm 25$  µM of Etoposide drug (SIGMA) was added and followed by incubation for 24 h. Cells were collected and fixed with 3.7% formaldehyde for 10 min at RT. Then cells were washed twice with PBS and permeabilized with 0.05% Triton X-100/PBS for 10 min at RT. Cells were treated with TUNEL<sup>®</sup> kit (TMR red-Roche) according to the manufacturer's protocol. Cells were then treated for FACS analysis using NOVOCYTE ACEA by exciting with 488 and 561 nm lasers at Montpellier Rio Imaging (MRI) facilities.

### **Transregulation experiments**

For luciferase assays,  $2.3 \times 10^5$  NIH 3T3 cells were seeded and transfected 24 h later, using JetPei method as described above, with 1 µg of the pAsLuc(Fire)-MLV-Luc(Reni) and different concentrations of p57cDNASD' (1.5 to 0.01 µg) or 1 µg of HA-P50 and derivative constructs. 48 h post-transfection, cells were washed with

cold PBS and then lysed in 1x passive lysis buffer (Promega), as described by the manufacturer. Luciferase assays were performed in a centro XS3 LB 960 microplate luminometer (Berthold Technologies) with respectively the Genofax A and Genofax C kit (Yelen). Luciferase activities were normalized with concentration of cell extract, determined by Bradford assay using a BSA standard set (Fermentas). For analysis of P50 effect on MLV promoter, NIH 3T3 were transfected, as described above, with F1/P50-GFP ratios of 1/2, 1/1 and 1/0.5. 48 h post-transfection, Gag and p50-GFP expressions were determined by Western blot analysis and the MLV FL RNA level was analyzed by RT-qPCR.

### **Chromatin immunoprecipitation assay**

A total of  $1.6 \times 10^6$  NIH 3T3 cells were co-transfected with 10 µg of plasmid expressing p1LTR-P50-GFP and 10 µg of F1 molecular clone using JetPEI in a 10 cm plate as described above. 48 h post-transfection, proteins from  $10^7$  cells were cross-linked to DNA with 3.7% formaldehyde for 10 min at RT. Cells were washed twice with PBS and harvested by scraping. Nuclei were isolated by incubation of cross-linked cells with 1 ml of Cell Fractionation Buffer Kit PARIS<sup>™</sup> (Invitrogen) for 5 min on ice and then pelleted by centrifugation (500xg). The nuclei were washed with 500 µl of Cell Fractionation Buffer Kit PARIS<sup>™</sup> (Invitrogen) and incubated with 500 µL of Nucleus Lysis Buffer for 3 h on rotator at 4°C. The chromatin was fragmented by adding 5 µl of MNase enzyme (SimpleChip<sup>®</sup>) at 37°C for 20 min to an average size of 200 to 500 bp. Cell debris were pelleted by centrifugation (20,000xg, 4°C), and supernatants were collected. Chromatin was diluted in 50 mM Tris-HCl, 167 mM NaCl, 1 mM EDTA, 0.01% SDS, 1.1% Triton X-100, protease inhibitor mix. To reduce nonspecific background, antibodies were preincubated with 15 µl Dynabeads and protein G (DYNAL<sup>®</sup>) according to the manufacturer's instructions. The antibody-bead complex was then added to chromatin samples, followed by incubation overnight at 4°C. Beads were washed once with low-salt buffer (2 mM Tris-HCl, 150 mM NaCl, 0.1% SDS, 1% Triton X-100, 2 mM EDTA), once with high-salt buffer (20 mM Tris-HCl, 500 mM NaCl, 0.1% SDS, 1% Triton X-100, 2 mM EDTA), once with LiCl wash buffer (10 mM Tris-HCl, 0.25 M LiCl, 1% Nonidet P-40, 1% sodium deoxycholate, 1 mM EDTA), and once with Tris-EDTA buffer. Chromatin was eluted from the beads in elution buffer (100 mM NaHCO<sub>3</sub>, 1% SDS) for 20 min at room temperature. DNA was incubated 5 h at 65°C and proteins were eliminated by a proteinase K digestion (1 h at 45°C). DNA was purified using the NucleoSpin<sup>®</sup> Gel and PCR clean-up kit and analyzed by PCR using One Taq<sup>®</sup> DNA Polymerase. To target the MLV LTR, the respective forward and reverse primers used were MLV sFM76:



5'GTGGTCTCGCTGTTCCCTTGGGA3' and MLV aM275 (5'GCAGGCGCATAAAATCAGTCATAG 3). As positive control we performed the PCR on molecular clone pMov9.1 and as a negative control, the PCR was performed on purified DNA from non-transfected cell extract.

### Transcriptomic analysis

#### *mRNA isolation and sequencing*

For each condition, three biological replicates were collected. Total RNA was extracted from cells with TriReagent (MRC) according to the manufacturer's instructions. RNA samples were treated with RQ1 DNase (Promega) in the presence of RNasin (Promega) for 25 min at 37 °C. RNA was extracted with phenol–chloroform then chloroform and finally precipitated with 100% ethanol. RNA pellets were washed with 70% ethanol and dissolved in water. Libraries of each replicate were constructed and sequencing (100 bp, paired-end) was performed using an Illumina HiSeq2000 sequencer by the Montpellier MGX sequencing platform (<http://www.mgx.cnrs.fr/>).

#### *Differential expression analysis*

Most bioinformatics analyses were performed by the MBB platform in Montpellier (<http://mbb.univ-montp2.fr/>). For each condition (CNT, CCI and CES) 3 biological replicates were sequenced to provide a mean of 50 million paired-end reads per sample. Pre-processing of the raw RNA-seq data to remove contaminants, adapters, low-quality sequences and unpaired reads was performed with TrimGalore (<https://github.com/FelixKrueger/TrimGalore>) with standard parameters (quality 30 and length 50 bp). The New Tuxedo protocol [95] was used for the assembly of transcripts and the quantification of gene expression levels before analyzing differential expression with DESeq2. First, reads from each sample were mapped to the mouse reference genome (GRCm38) with HISAT2-2.1.0 [96] with default parameters and `-dta` to provide *alignments tailored for transcript assemblers*. To assemble transcripts following this protocol, StringTie-1.3.3b was provided with `Mus_musculus.GRCm38.90.gtf` as a reference annotation gene model file downloaded from ensembl web site ([ftp://ftp.ensembl.org/pub/release-75/gtf/mus\\_musculus/](ftp://ftp.ensembl.org/pub/release-75/gtf/mus_musculus/)). This guided assembly was provided a parameter to limit the processing of read alignments to only the assembled transcripts matching the reference transcripts. The StringTie's accompanying script (<http://www.ccb.jhu.edu/software/stringtie/dl/prepDE.py>) was then used. The dendrogram representing the result of the clustering of all pairwise dissimilarities between the transcript expressions was obtained with the help of the Trinity tool (<https://github.com/trinityrnaseq/trinityrnaseq/wiki/Trinity-Differential-Expression>).

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12977-023-00631-w>.

**Additional file 1:** Table S1. Cellular RNAs up-regulated by the P50 MLV protein.

**Additional file 2:** Table S2. Cellular RNAs down-regulated by the P50 MLV protein.

### Acknowledgements

We are grateful to the Montpellier MRI imaging staff for technical assistance for microscopy and FACS. We thank Dr. Masaaki Miyazawa, for the gifts of the H187 antibody. Many thanks to Dr. R. Kiernan for critical reading of the manuscript.

### Authors' contributions

C.A.: Methodology, software, Formal analysis, writing—original draft preparation; J.F.: Methodology; F.L.D.: Microscopy and analysis. K.B.: Methodology, software, formal analysis; N.G.: Methodology; Resources; J.M.P.: Methodology, software, formal analysis; M.M.: Conceptualization, methodology, software, validation, investigation, funding acquisition; S.L.: Conceptualization, methodology, software, validation, investigation, writing—original draft preparation, funding acquisition, project administration and supervision. All authors participated in reviewing, editing, and reading and approved the final manuscript. All authors have read and agreed to the published version of the manuscript.

### Funding

This study was partially supported by the Ligue contre le cancer de l'Hérault, the CNRS and the Université de Montpellier. C.A. and F.L.D. received fellowships from the Palestinian Students Fund (PSF) and infectiopole sud, respectively.

### Data availability

All data generated or analyzed during this study are included in this published article and its Additional files 1 and 2.

### Declarations

#### Competing interests

The authors declare no competing interests.

#### Ethics approval and consent to participate

Consent was obtained from all study participants at the time of recruitment.

#### Consent for publication

Not applicable.

Received: 10 April 2023 / Accepted: 18 August 2023

Published online: 12 September 2023

### References

1. Tram J, Mesnard J-M, Peloponese J-M. Alternative RNA splicing in cancer: what about adult T-cell leukemia? *Front Immunol.* 2022;13:959382.
2. Martin GS. The hunting of the src. *Nat Rev Mol Cell Biol.* 2001;2:467–75.
3. Fan H. Leukemogenesis by Moloney murine leukemia virus: a multistep process. *Trends Microbiol.* 1997;5:74–82.
4. De Ravin SS, Su L, Theobald N, Choi U, Macpherson JL, Poidinger M et al. Enhancers Are Major Targets for Murine Leukemia Virus Vector Integration. Beemon KL, editor. *J Virol.* 2014;88:4504–13.
5. LaFave MC, Varshney GK, Gildea DE, Wolfsberg TG, Baxevanis AD, Burgess SM. MLV integration site selection is driven by strong enhancers and active promoters. *Nucleic Acids Res.* 2014;42:4257–69.

6. Kim R, Trubetskoy A, Suzuki T, Jenkins NA, Copeland NG, Lenz J. Genome-based identification of Cancer genes by Proviral Tagging in Mouse Retrovirus-Induced T-Cell Lymphomas. *J Virol*. 2003;77:2056–62.
7. Li J, Shen H, Himmel KL, Dupuy AJ, Largaespada DA, Nakamura T, et al. Leukaemia disease genes: large-scale cloning and pathway predictions. *Nat Genet*. 1999;23:348–53.
8. Lund AH, Turner G, Trubetskoy A, Verhoeven E, Wientjens E, Hulsman D, et al. Genome-wide retroviral insertional tagging of genes involved in cancer in Cdkn2a-deficient mice. *Nat Genet*. 2002;32:160–5.
9. Mikkers H, Allen J, Knipscheer P, Romeyn L, Hart A, Vink E, et al. High-throughput retroviral tagging to identify components of specific signaling pathways in cancer. *Nat Genet*. 2002;32:153–9.
10. Suzuki T, Shen H, Akagi K, Morse HC, Malley JD, Naiman DQ, et al. New genes involved in cancer identified by retroviral tagging. *Nat Genet*. 2002;32:166–74.
11. Déjardin J, Bompard-Maréchal G, Audit M, Hope TJ, Sitbon M, Mougél M. A Novel Subgenomic Murine Leukemia Virus RNA transcript results from alternative splicing. *J Virol*. 2000;74:3709–14.
12. Houzet L. A new retroelement constituted by a natural alternatively spliced RNA of murine replication-competent retroviruses. *EMBO J*. 2003;22:4866–75.
13. Maurel S, Houzet L, Garcia EL, Telesnitsky A, Mougél M. Characterization of a natural heterodimer between MLV genomic RNA and the SD' retroelement generated by alternative splicing. *RNA*. 2007;13:2266–76.
14. Maurel S, Mougél M. Murine leukemia virus RNA dimerization is coupled to transcription and splicing processes. *Retrovirology*. 2010;7:64.
15. Audit M, Déjardin J, Hohl B, Sidobre C, Hope TJ, Mougél M, et al. Introduction of a cis-acting mutation in the Capsid-Coding Gene of Moloney Murine Leukemia Virus extends its Leukemogenic Properties. *J Virol*. 1999;73:10472–9.
16. Ramirez JM, Houzet L, Koller R, Bies J, Wolff L, Mougél M. Activation of c-myc by 5' retrovirus promoter insertion in myeloid neoplasms is dependent upon an intact alternative splice donor site (SD') in gag. *Virology*. 2004;330:398–407.
17. Evans LH, Dresler S, Kabat D. Synthesis and glycosylation of Polyprotein Precursors to the Internal Core Proteins of Friend Murine Leukemia Virus. *J Virol*. 1977;24:865–74.
18. Edwards SA, Fan H. Gag-related polyproteins of Moloney murine leukemia virus: evidence for independent synthesis of glycosylated and unglycosylated forms. *J Virol*. 1979;30:551–63.
19. Saris CJ, van Eenbergen J, Liskamp RM, Bloemers HP. Structure of glycosylated and unglycosylated gag and gag-pol precursor proteins of Moloney murine leukemia virus. *J Virol*. 1983;46:841–59.
20. Prats A-C, De Billy G, Wang P, Darlix J-L. CUG initiation codon used for the synthesis of a cell surface antigen coded by the murine leukemia virus. *J Mol Biol*. 1989;205:363–72.
21. Zhao W, Akkawi C, Mougél M, Ross SR. Murine Leukemia Virus P50 Protein Counteracts APOBEC3 by Blocking Its Packaging. Simon V, editor. *J Virol*. 2020;94:e00032-20.
22. Stavrou S, Nitta T, Kotla S, Ha D, Nagashima K, Rein AR, et al. Murine leukemia virus glycosylated gag blocks apolipoprotein B editing complex 3 and cytosolic sensor access to the reverse transcription complex. *Proc Natl Acad Sci USA*. 2013;110:9078–83.
23. Rosales Gerpe MC, Renner TM, Bélanger K, Lam C, Aydin H, Langlois M-A. N-Linked Glycosylation Protects Gammaretroviruses against Deamination by APOBEC3 Proteins. Ross SR, editor. *J Virol*. 2015;89:2342–57.
24. Renner TM, Bélanger K, Lam C, Gerpe MCR, McBane JE, Langlois M-A. Full-Length Glycosylated Gag of Murine Leukemia Virus Can Associate with the Viral Envelope as a Type I Integral Membrane Protein. Simon V, editor. *J Virol*. 2018;92:e01530-17.
25. Roth MJ, Schwartzberg P, Tanese N, Goff SP. Analysis of mutations in the integration function of Moloney murine leukemia virus: effects on DNA binding and cutting. *J Virol*. 1990;64:4709–17.
26. Aiyer S, Rossi P, Malani N, Schneider WM, Chandar A, Bushman FD, et al. Structural and sequencing analysis of local target DNA recognition by MLV integrase. *Nucleic Acids Res*. 2015;43:5647–63.
27. Wanaguru M, Barry DJ, Benton DJ, O'Reilly NJ, Bishop KN. Murine leukemia virus p12 tethers the capsid-containing pre-integration complex to chromatin by binding directly to host nucleosomes in mitosis. Beemon K, editor. *PLoS Pathog*. 2018;14:e1007117.
28. Yueh A, Goff SP. Phosphorylated serine residues and an arginine-rich domain of the Moloney Murine Leukemia Virus p12 protein are required for early events of viral infection. *J Virol*. 2003;77:1820–9.
29. Aiyer S, Swapna GVT, Malani N, Aramini JM, Schneider WM, Plumb MR, et al. Altering murine leukemia virus integration through disruption of the integrase and BET protein family interaction. *Nucleic Acids Res*. 2014;42:5917–28.
30. De Rijck J, de Kogel C, Demeulemeester J, Vets S, El Ashkar S, Malani N, et al. The BET family of proteins targets Moloney murine leukemia virus integration near transcription Start Sites. *Cell Rep*. 2013;5:886–94.
31. Mann R, Mulligan RC, Baltimore D. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell*. 1983;33:153–9.
32. Risser R, Pollack R. A nonselective analysis of SV40 transformation of mouse 3T3 cells. *Virology*. 1974;59:477–89.
33. Baluyot MF, Grosse SA, Lyddon TD, Janaka SK, Johnson MC. CRM1-Dependent trafficking of retroviral gag proteins revisited. *J Virol*. 2012;86:4696–700.
34. Arpin-André C, Laverdure S, Barbeau B, Gross A, Mesnard J-M. Construction of a reporter vector for analysis of bidirectional transcriptional activity of retrovirus LTR. *Plasmid*. 2014;74:45–51.
35. Rasmussen MH, Ballarín-González B, Liu J, Lassen LB, Füchtbauer A, Füchtbauer E-M, et al. Antisense transcription in Gammaretroviruses as a mechanism of Insertional activation of host genes. *J Virol*. 2010;84:3780–8.
36. Wang G, Liu H, An L, Hou S, Zhang Q. CAPG facilitates diffuse large B-cell lymphoma cell progression through PI3K/AKT signaling pathway. *Hum Immunol*. 2022;83:832–42.
37. Zhang Y, Frohman MA. Cellular and physiological roles for phospholipase D1 in Cancer. *J Biol Chem*. 2014;289:22567–74.
38. Burdelski C, Borchering L, Kluth M, Hube-Magg C, Melling N, Simon R, et al. Family with sequence similarity 13 C (FAM13C) overexpression is an independent prognostic marker in prostate cancer. *Oncotarget*. 2017;8:31494–508.
39. Suzuki Y, Kobayashi M, Miyashita H, Ohta H, Sonoda H, Sato Y. Isolation of a small vasohibin-binding protein (SVBP) and its role in vasohibin secretion. *J Cell Sci*. 2010;123:3094–101.
40. Liu Z, Lv YJ, Song YP, Li XH, Du YN, Wang CH, et al. Lysosomal membrane protein TMEM192 deficiency triggers crosstalk between autophagy and apoptosis in HepG2 hepatoma cells. *Oncol Rep*. 2012;28:985–91.
41. Zhou B, Irwanto A, Guo Y-M, Bei J-X, Wu Q, Chen G, et al. Exome sequencing and digital PCR analyses reveal novel mutated genes related to the metastasis of pancreatic ductal adenocarcinoma. *Cancer Biol Ther*. 2012;13:871–9.
42. Laurin M, Huber J, Pelletier A, Houalla T, Park M, Fukui Y, et al. Rac-specific guanine nucleotide exchange factor DOCK1 is a critical regulator of HER2-mediated breast cancer metastasis. *Proc Natl Acad Sci USA*. 2013;110:7434–9.
43. Varland S, Myklebust LM, Goksøyr S, Glomnes N, Torsvik J, Varhaug JE, et al. Identification of an alternatively spliced nuclear isoform of human N-terminal acetyltransferase Naa30. *Gene*. 2018;644:27–37.
44. Ge W, Jiang M, Zhang F, Ma Y, Wang H, Xu Y. ZGRF1 is Associated with Poor Prognosis in Triple-Negative breast Cancer and promotes Cancer Stemness based on Bioinformatics. *OTT*. 2020;13:2843–54.
45. Peng Y, Li M, Clarkson BD, Pehar M, Lao PJ, Hillmer AT, et al. Deficient import of Acetyl-CoA into the ER lumen causes neurodegeneration and propensity to infections, inflammation, and Cancer. *J Neurosci*. 2014;34:6772–89.
46. Loaeza-Loaeza J, Cerecedo-Castillo AJ, Rodríguez-Ruiz HA, Castro-Coronel Y, Del Moral-Hernández O, Recillas-Targa F, et al. DNMT3B overexpression downregulates genes with CpG islands, common motifs, and transcription factor binding sites that interact with DNMT3B. *Sci Rep*. 2022;12:20839.
47. Vater I, Wagner F, Kreuz M, Berger H, Martin-Subero JI, Pott C, et al. GeneChip analyses point to novel pathogenetic mechanisms in mantle cell lymphoma. *Br J Haematol*. 2009;144:317–31.
48. Zhang J, Huang JY, Chen YN, Yuan F, Zhang H, Yan FH, et al. Whole genome and transcriptome sequencing of matched primary and peritoneal metastatic gastric carcinoma. *Sci Rep*. 2015;5:13750.
49. Colas E, Perez C, Cabrera S, Pedrola N, Monge M, Castellvi J, et al. Molecular markers of endometrial carcinoma detected in uterine aspirates. *Int J Cancer*. 2011;129:2435–44.
50. Neubert NJ, Soneson C, Barras D, Baumgaertner P, Rimoldi D, Delorenzi M, et al. A well-controlled experimental system to study interactions of cytotoxic T lymphocytes with Tumor cells. *Front Immunol*. 2016;7:326.
51. Bian K, Muppani NR, Elkhadrady L, Wang W, Zhang C, Chen T, et al. ERK3 regulates TDP2-mediated DNA damage response and chemoresistance in lung cancer cells. *Oncotarget*. 2016;7:6665–75.
52. Melville\*\* MW, Katze\* MG, Tan S-L. P58IPK, a novel cochaperone containing tetratricopeptide repeats and a J-domain with oncogenic potential: CMLS. *Cell Mol Life Sci*. 2000;57:311–22.
53. Celik H, Koh WK, Kramer AC, Ostrander EL, Mallaney C, Fisher DAC, et al. JARID2 functions as a tumor suppressor in myeloid neoplasms by

- repressing Self-Renewal in hematopoietic progenitor cells. *Cancer Cell*. 2018;34:741–756e8.
54. Badie C, Blachowicz A, Barjaktarovic Z, Finnon R, Michaux A, Sarioglu H, et al. Transcriptomic and proteomic analysis of mouse radiation-induced acute myeloid leukaemia (AML). *Oncotarget*. 2016;7:40461–80.
55. Agashe RP, Lippman SM, Kurzrock R. JAK: not just another kinase. *Mol Cancer Ther*. 2022;21:1757–64.
56. Jang J-E, Kim H-P, Han S-W, Jang H, Lee S-H, Song S-H, et al. NFATC3–PLA2G15 Fusion transcript identified by RNA sequencing promotes Tumor Invasion and Proliferation in Colorectal Cancer Cell Lines. *Cancer Res Treat*. 2019;51:391–401.
57. Corveleyn A, Janssen H, Martini A, Somers R, Cools J, Marynen P. Cellular transformation of NIH3T3 fibroblasts by Cliz/NMP4 fusions. *J Cell Biochem*. 2005;94:1112–25.
58. Zhang YJ, Zhong XP, Chen Y, Liu SR, Wu G, Liu YF. Association between CASP-8 gene polymorphisms and cancer risk in some asian population based on a HuGE review and meta-analysis. *Genet Mol Res*. 2013;12:6466–76.
59. Huber FM, Greenblatt SM, Davenport AM, Martinez C, Xu Y, Vu LP, et al. Histone-binding of DPf2 mediates its repressive role in myeloid differentiation. *Proc Natl Acad Sci USA*. 2017;114:6016–21.
60. Dubois N, Marquet R, Paillart J-C, Bernacchi S. Retroviral RNA dimerization: from structure to functions. *Front Microbiol*. 2018;9:527.
61. Didierlaurent L, Racine PJ, Houzet L, Chamontin C, Berkhout B, Mougel M. Role of HIV-1 RNA and protein determinants for the selective packaging of spliced and unspliced viral RNA and host U6 and 7SL RNA in virus particles. *Nucleic Acids Res*. 2011;39:8915–27.
62. Ensofi B, Barillari G, Salahuddin SZ, Gallo RC, Wong-Staal F. Tat protein of HIV-1 stimulates growth of cells derived from Kaposi's sarcoma lesions of AIDS patients. *Nature*. 1990;345:84–6.
63. Musinova YR, Sheval EV, Dib C, Germini D, Vassetzky YS. Functional roles of HIV-1 Tat protein in the nucleus. *Cell Mol Life Sci*. 2016;73:589–601.
64. Hanahan D, Weinberg RA. Hallmarks of Cancer: the Next Generation. *Cell*. 2011;144:646–74.
65. Watson K, Edwards RJ. HIV-1 trans-activating (Tat) protein. *Biochem Pharmacol*. 1999;58:1521–8.
66. Takachi T, Takahashi M, Takahashi-Yoshita M, Higuchi M, Obata M, Mishima Y, et al. Human T-cell leukemia virus type 1 Tax oncoprotein represses the expression of the BCL 11 B tumor suppressor in T-cells. *Cancer Sci*. 2015;106:461–5.
67. Ali A, Mishra R, Kaur H, Chandra Banerjee A. HIV-1 Tat: an update on transcriptional and non-transcriptional functions. *Biochimie*. 2021;190:24–35.
68. Nyborg JK, Egan D, Sharma N. The HTLV-1 tax protein: revealing mechanisms of transcriptional activation through histone acetylation and nucleosome disassembly. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. 2010;1799:266–74.
69. Sharma A, Larue RC, Plumb MR, Malani N, Male F, Slaughter A, et al. BET proteins promote efficient murine leukemia virus integration at transcription start sites. *Proc Natl Acad Sci USA*. 2013;110:12036–41.
70. Winans S, Goff SP. Mutations altering acetylated residues in the CTD of HIV-1 integrase cause defects in proviral transcription at early times after integration of viral DNA. Krausslich H-G, editor. *PLoS Pathog*. 2020;16:e1009147.
71. Rocchi C, Louvat C, Miele AE, Batisse J, Guillon C, Ballut L, et al. The HIV-1 integrase C-Terminal domain induces TAR RNA structural changes promoting Tat binding. *IJMS*. 2022;23:13742.
72. Wight DJ, Boucherit VC, Nader M, Allen DJ, Taylor IA, Bishop KN. The Gamaretroviral p12 protein has multiple domains that function during the early stages of replication. *Retrovirology*. 2012;9:83.
73. Curren R, Van Duyn R, Jaworski E, Guendel I, Sampey G, Das R, et al. HTLV Tax: a fascinating multifunctional co-regulator of viral and Cellular Pathways. *Front Microbiol*. 2012;3:1–23.
74. He H, Dai F, Yu L, She X, Zhao Y, Jiang J, et al. Identification and characterization of nine Novel Human Small GTPases showing variable expressions in Liver Cancer Tissues. *gene expr*. 2002;10:231–42.
75. Jeong SW, Jang JY, Chung RT. Hepatitis C virus and hepatocarcinogenesis. *Clin Mol Hepatol*. 2012;18:347.
76. Mishra J, Das JK, Kumar N. Janus kinase 3 regulates adherens junctions and epithelial mesenchymal transition through  $\beta$ -catenin. *J Biol Chem*. 2017;292:16406–19.
77. Jiang Z, Lao T, Qiu W, Polverino F, Gupta K, Guo F, et al. A chronic obstructive Pulmonary Disease susceptibility gene, *FAM13A*, regulates Protein Stability of  $\beta$ -Catenin. *Am J Respir Crit Care Med*. 2016;194:185–97.
78. Yin S-Q, Wang J-J, Zhang C-M, Liu Z-P. The development of MetAP-2 inhibitors in Cancer Treatment. *CMC*. 2012;19:1021–35.
79. Dahlke C, Maul K, Christalla T, Walz N, Schult P, Stocking C et al. JSimaseditor2012A microRNA encoded by Kaposi Sarcoma-Associated Herpesvirus promotes B-Cell expansion in vivo. *PLoS ONE* e49435.
80. Zemel R, Issachar A, Tur-Kaspa R. The role of oncogenic viruses in the pathogenesis of Hepatocellular Carcinoma. *Clin Liver Dis*. 2011;15:261–79.
81. Nie K, Li J, He X, Wang Y, Zhao Q, Du M, et al. COX6B2 drives metabolic reprogramming toward oxidative phosphorylation to promote metastasis in pancreatic ductal cancer cells. *Oncogenesis*. 2020;9:51.
82. Vidoni S, Harbour ME, Guerrero-Castillo S, Signes A, Ding S, Fearnley IM, et al. MR-15 interacts with PET100 and PET117 in Module-Based Assembly of Human cytochrome c oxidase. *Cell Rep*. 2017;18:1727–38.
83. Buettner C, Muse ED, Cheng A, Chen L, Scherer T, Poca A, et al. Leptin controls adipose tissue lipogenesis via central, STAT3-independent mechanisms. *Nat Med*. 2008;14:667–75.
84. De Vivo DC, DiMauro S. Mitochondrial Diseases. *Swaiman's Pediatric Neurology*. Elsevier; 2017. pp. 334–46.
85. Alkhateeb AM, Aburahma SK, Habbab W, Thompson IR. Novel mutations in *WFOX*, *RARS2*, and *C10orf2* genes in consanguineous arab families with intellectual disability. *Metab Brain Dis*. 2016;31:901–7.
86. Roth E, Frohman MA. Proliferative and metastatic roles for phospholipase D in mouse models of cancer. *Adv Biol Regul*. 2018;67:134–40.
87. Du H, Zhao J, Hai L, Wu J, Yi H, Shi Y. The roles of vasohibin and its family members: beyond angiogenesis modulators. *Cancer Biol Ther*. 2017;18:827–32.
88. Fu M, Blackshear PJ. RNA-binding proteins in immune regulation: a focus on CCCH zinc finger proteins. *Nat Rev Immunol*. 2017;17:130–43.
89. Fridman WH, Sautès-Fridman C. Le microenvironnement tumoral: Matrice nourricière, champ de bataille et cible thérapeutique des cancers. *Med Sci (Paris)*. 2014;30:359–65.
90. Mesri EA, Feitelson MA, Munger K. Human viral oncogenesis: a Cancer Hallmarks Analysis. *Cell Host Microbe*. 2014;15:266–82.
91. Chopard C, Tong PBV, Tóth P, Schatz M, Yezid H, Debaisieux S, et al. Cyclophilin A enables specific HIV-1 Tat palmitoylation and accumulation in uninfected cells. *Nat Commun*. 2018;9:2251.
92. Mougel M, Akkawi C, Chamontin C, Feuillard J, Pessel-Vivares L, Socol M, et al. NXF1 and CRM1 nuclear export pathways orchestrate nuclear export, translation and packaging of murine leukaemia retrovirus unspliced RNA. *RNA Biol*. 2020;17:528–38.
93. Chamontin C, Yu B, Racine P-J, Darlix J-L, Mougel M. MoMuLV and HIV-1 Nucleocapsid Proteins Have a Common Role in Genomic RNA Packaging but Different in Late Reverse Transcription. Pöhlmann S, editor. *PLoS ONE*. 2012;7:e51534.
94. Ke N, Albers A, Claassen G, Yu D, Chatterton JE, Hu X, et al. One-week 96-well soft agar growth assay for cancer target validation. *Biotechniques*. 2004;36:826–33.
95. Perteau M, Kim D, Perteau GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc*. 2016;11:1650–67.
96. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12:357–60.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.