I.

OPEN

LINE-1 ORF1p is a Promising Biomarker in Cervical Intraepithelial Neoplasia Degree Assessment

Réka Karkas, MD,† Khaldoon Sadiq Ahmed Abdullah, MSc,*† László Kaizer, MD, FRCPath,‡ Ádám Ürmös, BA,§ May Raya, MSc,*† Lilla Tiszlavicz, MD,‖ Tibor Pankotai, PhD,‡§¶ István Nagy, PhD,#** Lajos Mátés, PhD,* and Farkas Sükösd, MD, PhD‡*

Abstract: Cervical intraepithelial neoplasia (CIN) represents a spectrum of preinvasive squamous lesions within the cervical epithelium, whose identification is a diagnostic challenge due to subtle histomorphological differences among its categories. This study explores ORF1p, a nucleic acid-binding protein derived from long interspersed nuclear element-1 (LINE-1), as a potential biomarker for enhancing CIN diagnosis. A comprehensive analysis of 143 cervical specimens, encompassing CIN I (n = 20), CIN II (n = 46), CIN III (n = 14), invasive cancer (n = 32), and nondysplastic cases (normal cervical epithelia (n = 24) and atrophy (n = 7) were conducted. ORF1p, Ki67, and p16 expressions were evaluated using immunohistochemistry. ORF1p immunopositivity was detected in the vast majority [110/112 (98.2%)] of dysplastic and neoplastic (CIN and invasive cancer) specimens, whereas 19/24 (79.2%) of normal cervical specimens lacked ORF1p expression. The observed pattern of ORF1p expression showed a progressively increasing extent and intensity with advancing CIN grades. CIN I exhibited mild ORF1p expression in the lower one or two-thirds of the cervical epithelium [14/16 (87.5%)], whereas CIN II demonstrated moderate to strong ORF1p expression spanning the lower two-thirds [29/46 (63.0%)]. Pronounced transepithelial ORF1p immunopositivity characterized CIN III cases [13/14 (92.8%)] and cervical cancer [30/32 (93.8%)]. These findings propose ORF1p as a valuable indicator even for detecting CIN I, effectively discerning them from normal cervical tissue ($p < 0.0001$). Our findings underscore the potential of ORF1p as an early diagnostic marker for cervical neoplasia.

Key Words: cervical intraepithelial neoplasia, diagnostic biomarker, ORF1p, LINE-1, carcinogenesis

(*Int J Gynecol Pathol* 2025;44:22–30)

From the *Laboratory of Cancer Genome Research, Institute of Genetics, HUN-REN Biological Research Centre, Szeged, Hungary; †Doctoral School of Multidisciplinary Medical Sciences, University of Szeged, Albert Szent-Györgyi Medical School, Szeged, Hungary; ‡Department of Pathology, Albert Szent-Györgyi Health Centre, University of Szeged, Szeged, Hungary; §Genome Integrity and DNA Repair Core Group, Hungarian Centre of Excellence for Molecular Medicine, Szeged, Hungary; ‖Department of Pediatrics and Pediatric Health Centre, Albert Szent-Györgyi Health Centre, University of Szeged, Szeged, Hungary; ¶Competence Centre of the Life Sciences Cluster of the Centre of Excellence for Interdisciplinary Research, Development and Innovation, University of Szeged, Hungary; #Seqomics Biotechnology Ltd, Mórahalom, Hungary; and **Sequencing Platform, Institute of Biochemistry, HUN-REN Biological Research Centre, Szeged, Hungary.

The project was, in part, funded by the 2021-1.1.4- GYORSÍTÓSÁV-2022-00018 project of the National Research, Development and Innovation Fund, Hungary. R.K. was supported by the ÚNKP-23-3 New National Excellence Program of the Ministry for Culture and Innovation from the source of the National Research, Development, and Innovation Fund, Hungary.

R.K. and K.S.A.A. are joint first authors.

The authors declare no conflict of interest.

Address correspondence to Farkas Sükösd, MD, PhD, Department of Pathology, University of Szeged, Állomás Street 1, Szeged H-6721, Hungary. E-mail: sukosd.farkas@med.u-szeged.hu, and Lajos Mátés, PhD, Laboratory of Cancer Genome Research, Institute of Genetics, HUN-REN Biological Research Centre, Temesvári Blvd. 62, Szeged H-6726, Hungary. E-mail: mates.lajos@brc.hu.

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.intjgynpathology.com.

DOI: 10.1097/PGP.0000000000001035

Cervical cancer is a significant global health burden and a leading cause of cancer-related morbidity and mortality among women.[1] It is the fourth most common cancer in women worldwide,[2] primarily caused by persistent infection with high-risk human papillomavirus (HPV).[3,4] Despite advances in prevention and early detection, cervical cancer continues to pose a substantial threat, particularly in low and middle-income countries with limited access to screening programs and HPV vaccination.[5] Understanding the importance of early detection and timely intervention as key preventive measures can significantly reduce the burden of this disease and improve women's health outcomes worldwide.[6]

Preinvasive squamous lesions of the cervix, collectively known as cervical intraepithelial neoplasia (CIN), encompass a range of cellular changes that occur within the cervical epithelium. CIN is commonly classified into different grades based on the severity of the observed cellular abnormalities. The CIN I is a low-grade squamous lesion (LSIL) mostly caused by low-risk HPV types that regress spontaneously in more than three-quarters of cases. By contrast, high-grade squamous cell lesions (HSILs), such as

CIN II and CIN III, have a significantly higher risk of progression to invasive cervical cancer. This difference highlights the clinical significance of accurate diagnosis.[6]

The diagnosis of CIN primarily relies on histological evaluation of cervical biopsy specimens. However, interobserver variability and subjectivity in interpreting histopathological features of CIN lesions can impact diagnostic accuracy and reproducibility.[7–10] Therefore, there is a growing need for objective and reliable biomarkers that can complement traditional diagnostic approaches and improve the accuracy of CIN diagnosis. The most commonly used immunohistochemical markers are Ki67 and p16.[11] Ki67 is a marker of cellular proliferation dominantly present at sites of ribosomal RNA transcription.[12] In the context of cervical lesions, a high Ki67 labeling index (with more than 30% of the cells exhibiting nuclear immunopositivity) is associated with more aggressive or advanced lesions, such as high-grade CIN or cervical cancer.[13] The p16, also known as p16INK4a, is a tumor suppressor protein that acts as a negative regulator of the cell cycle, inhibiting the phosphorylation of retinoblastoma (Rb) protein and preventing cell proliferation. In high-risk HPV-associated lesions, the viral oncoproteins E6 and E7 disrupt the normal function of p53 and Rb, respectively, leading to increased p16 expression as a cellular response to this dysregulation.[14] Consequently, cytoplasmic p16 overexpression has been recognized as a surrogate marker for high-risk HPV infection and a reliable indicator of HPV-associated neoplasia.[15,16] In the cases of HSIL and carcinoma in situ, HPV genotypes 16, 31, 33, and 35 are primarily involved. Consequently, these cases exhibit p16 immunopositivity.[4] Conversely, CIN I lesions are associated with a more diverse range of over 20 different HPV genotypes and are not expected to express p16. However, in 25% to 30% of LSIL cases, high-risk HPV infection can be identified, rendering their positive p16 status an insufficient classifier of the degree of CIN.[17,18] This incorrect application of p16 IHC can overdiagnose HSIL when performed on unequivocal low-grade lesions or when it upgrades questionable lesions based on non-block p16 staining patterns.[19] According to the follow-up studies, p16 immunopositivity does not correlate with LSIL to HSIL progression either,[17] so it is only the negative p16 staining result that has diagnostic value, suggesting a case of LSIL.

Long Interspersed Nuclear Elements 1 (LINE-1) are autonomous retrotransposons accounting for 17% of the genome, where their copy number reaches 500,000.[20] LINE-1 retrotransposons contribute to evolution and genomic diversity,[21–23] although they might have harmful genotoxic effects as well. LINE-1 elements can insert themselves into new genomic loci with a "copy-and-paste" mechanism by which they potentially disrupt the function of genes, contributing to the development of a variety of genetic disorders, including cancer.[24] The seeming conflict between their positive contribution to human evolution and their disease-causing genotoxicity is clarified by the fact that, under nonpathological conditions, their well-controlled activity is restricted to germ cells[25] and early embryonic development.[26] De novo retrotransposition events in the germline contribute to heritable genetic changes and the rate of new germline events was estimated as ~1 insertion for every 200 births based on genome comparisons.[27] The vast majority of these events have no discernible effect, but a very small proportion can cause familial inherited diseases.[24] By contrast, in differentiated tissues under nonpathological conditions, only sporadic and weak LINE-1 activity is observed in some epithelial cells and neurons.[28] However, when LINE-1 activity is initiated in somatic tissues, it leads to the development of cancers through the formation of cancer driver mutations.[24,29,30] In differentiated tissues, numerous defense mechanisms tightly suppress LINE-1 retrotransposition to maintain the integrity of the somatic genome.[31,32] However, under pathological conditions, these protective systems may decline, allowing the activation of LINE-1 elements, which in turn contribute to tumor development.

Active LINE-1 loci produce transcripts encoding 2 proteins, ORF1p and ORF2p. ORF1p acts as a high-affinity RNA binding protein and nucleic acid chaperone, playing a critical role in facilitating retrotransposition by binding to LINE-1 RNA and forming stable ribonucleoprotein particles.[33] ORF2p is responsible for reverse transcription and integration of LINE-1 elements.[34]

The emerging potential of ORF1p as a biomarker has been recognized in different types of cancer and their precursor lesions.[35] ORF1p expression serves as a characteristic feature in numerous cancer types, such as breast, ovarian, bladder, esophageal, and pancreatic cancer, among many others.[36–41] However, its association with CIN and cervical cancer has not been investigated.

This study aims to explore the expression of ORF1p in preneoplastic and neoplastic cervical specimens and evaluates its potential as a diagnostic biomarker for CIN. By investigating the diagnostic value of ORF1p, we aim to enhance the accuracy and objectivity of CIN diagnosis. Identifying a reliable biomarker like ORF1p could aid in risk stratification and reduce interobserver variability in CIN diagnosis.

MATERIALS AND METHODS

Tissue Specimens

Archived pathological samples collected for diagnostic purposes were used for TMA assembly. The study was performed on 112 selected formalin-fixed paraffin-embedded (FFPE) specimens of CIN I (n = 20), CIN II (n = 46), CIN III (n = 14), and cervical cancer (n = 32). The control group comprised 31 cases of nondysplastic cervical tissue obtained from patients who underwent total hysterectomy for uterine leiomyomas and uterine prolapsus. These specimens were further categorized into the following groups: normal (n = 24) and atrophic epithelium (n = 7). Cores were chosen from each case, and their representativeness was reviewed after H&E staining. For tumor specimens, the following parameters were investigated: tumor grade, primary tumor size, lymph

node, and metastasis status (pTNM). All the patient data acquired in this study were anonymized and did not impact diagnosis or treatment. The age range of the patient cohort encompassed a spectrum of 22 to 83 years, with a mean age of 52 years, selected from the period between 2018 and 2022 at the Clinical Centre of Albert Szent-Györgyi Medical School at the University of Szeged, Hungary. (Supplementary Table 1, Supplemental Digital Content 1, http://links.lww.com/IJGP/A176) The research protocol was reviewed and approved by the Institutional Committee of Science and Research Ethics of the Medical Research Council, Budapest, Hungary (Reference number: BM/3049/2023).

Immunohistochemistry

Immunohistochemical staining was run on the Bond Max Autostainer (Leica Biosystems, Wetzlar, Germany) with Bond Polymer Detection System (Vision BioSystems, Newcastle upon Tyne, UK). Monoclonal mouse anti-p16 antibody (clone MX007, MAD-000690QD-12, Master Diagnostica S.L, Granada, Spain), monoclonal anti-Ki67 (clone SP6, #10080, Histopathology Ltd., Pécs, Hungary), and monoclonal anti-LINE-1 ORF1p antibody (clone 4H1, MABC1152, Millipore Darmstadt, Germany) were used. Negative controls were performed on all sections using an equivalent concentration of a subclass-matched IgG_{1K}. The IHC sections were scored by 2 experienced pathologists (L.K. and F.S.) independently to ensure interobserver agreement. The observers were blinded to any clinical parameter. A consensus outcome was reached in case of discordance.

To evaluate ORF1p immunostaining results, a modified version of the immunoreactive scale (IRS) by Remmele and Stegner was used.[42] The samples underwent scoring based on both staining intensity and the extent of epithelial involvement. The cytoplasmic staining intensity was assessed using a semi-quantitative scale: 0 (negative), 1 (weak), 2 (moderate), and 3 (strong). The extent of positively stained epithelium was scored as follows: 0 (absent), 1 (involving the basal one-third), 2 (involving the basal two-thirds), and 3 (involving the entire epithelial thickness). To determine the ORF1p immunoreaction score, the intensity and extent scores were multiplied (Table 1). The ORF1p staining was graded negative in the presence of less than 10% weakly reactive epithelial cells.[43]

The degree of p16 expression was quantified based on the percentage of p16-positive cells. In a semi-quantitative scale, a 0 designation was assigned when the proportion of positive cells fell below 1%. Grades 1 and 2 were allocated to cases with clustered positive cells and percentages of positive cells ranging from 1% to 5% and 5% to 25%, respectively. Grade 3 was designated in instances where widespread positive cells constituted more than 25% of the total.[13,44]

To assess the degree of Ki67 expression, the nuclei of 200 epithelial cells spanning the entire epithelial layer were examined in each specimen. The Ki67 index was established as the percentage of Ki67 positive cells. Grades 1, 2, and 3 were applied to categorize the extent of expression. Grade 1, 2, and 3 were given when the Ki67 index was below 5%, 5% to 30%, and greater than 30%, respectively.[13,45]

TABLE 1. Modified version of the immunoreactive scale (IRS) scoring system[42] for the evaluation of ORF1p immunohistochemistry in cervical specimens. (A) The measured parameters and their scaling (B) The final immunoreaction scores and their classification

A

Positively stained epithelial thickness (PP)	PP score	Intensity of staining (SI)	SI score
No staining	0	No color reaction	0
Basal one-third	1	Mild reaction	1
Basal two-thirds	2	Moderate reaction	2
Transepithelial	3	Strong reaction	3

B

IRS Scores (PP × SI)	IRS Classification
0	Negative
1–2	Positive, mild immunoreaction
3–4	Positive, moderate immunoreaction
6–9	Positive, strong immunoreaction

Statistical Analysis

Staining score data handling and analysis were carried out in the R statistical programming environment.[46] Comparisons across entities were made using Pearson χ^2 and Fisher exact tests. The associations were considered statistically significant for $P < 0.05$ in all analyzed cases.

RESULTS

We investigated the LINE-1 ORF1p immunopositivity of 112 formalin-fixed paraffin-embedded (FFPE) specimens of CIN I (n = 20), CIN II (n = 46), CIN III (n = 14), and cervical cancer (n = 32), and a control group of 31 nondysplastic cervical tissue samples. The ORF1p exhibited increasing expression with increasing grades of CIN towards cancer, while it was barely or undetectable in normal cervical epithelium (Fig. 1). We also performed immunostaining for the protein markers Ki67 and p16 in our entire sample panel. The Ki67 expression deviates from the normal pattern in the CIN II stage only (Fig. 1). Similarly, abnormal expression of the p16 marker is usually not observed before CIN II stage (Fig. 1). In contrast, the expression of ORF1p typically shows an abnormal pattern already from CIN I stage onwards, with a characteristic increase in immunopositivity with increasing CIN grading (Fig. 1). Representative images of CIN I stage at higher magnification clearly show patterns of Ki67 and p16 immunostaining typical of normal cervical epithelium, while ORF1p overexpression is easily discernible from normal staining patterns (Fig. 2).

The ORF1p positivity often does not extend to the full thickness of the cervical epithelia. In some cases, only the lower one-third, in other cases, the lower two-thirds,

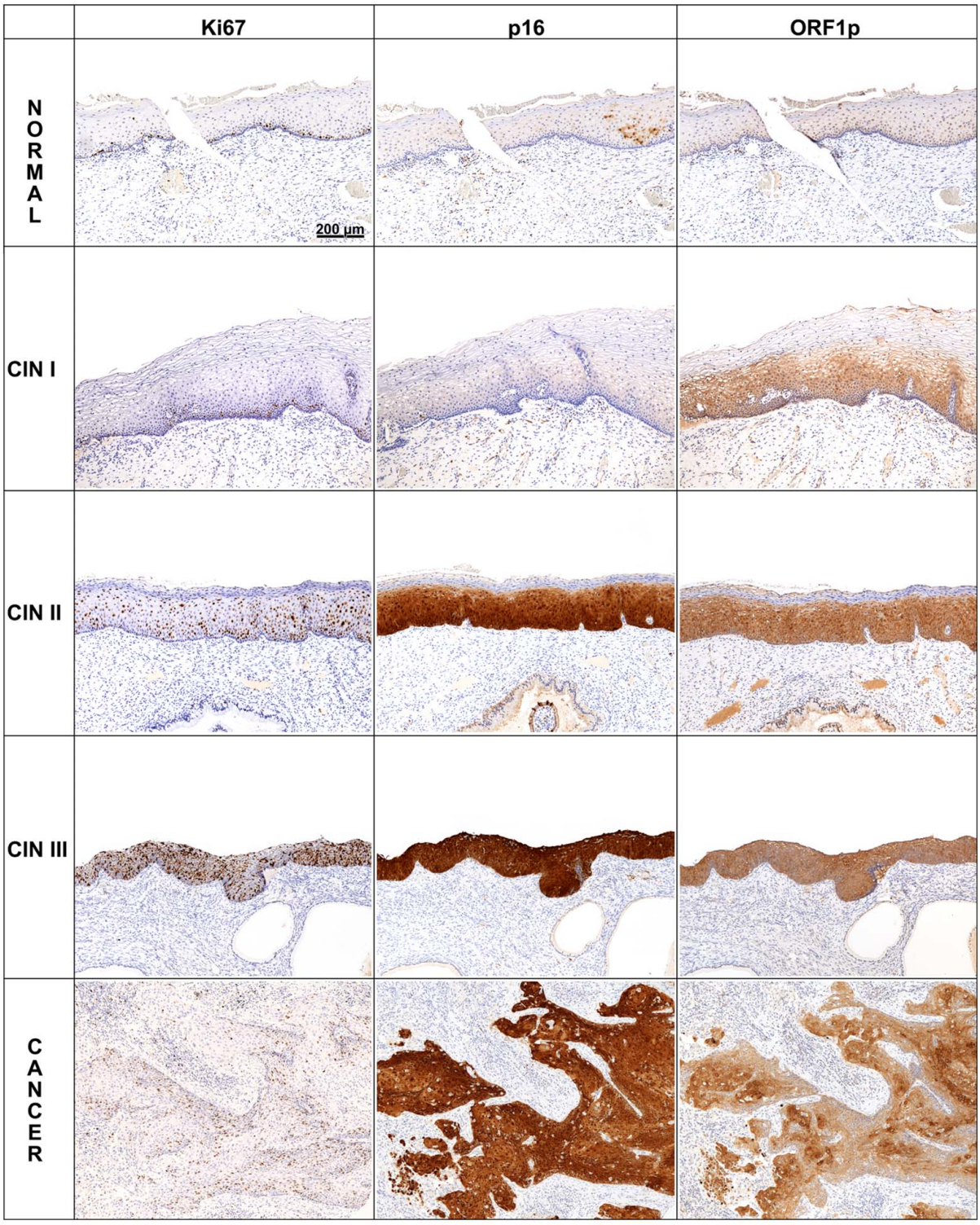

FIGURE 1. Immunostaining for ORF1p, p16 and Ki67 across differently graded lesions of cervical epithelium progressing towards cancer. Representative sections of FFPE specimens from normal cervical epithelium, cervical epithelial hyperplasia, cervical epithelial neoplasia (CIN I, CIN II, and CIN III) and cancer are presented. Scale bar = 200 µm.

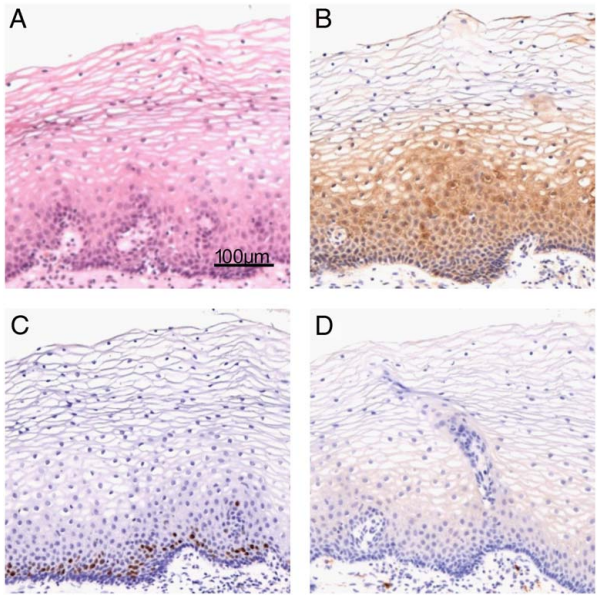

FIGURE 2. Immunostaining for ORF1p, p16, and Ki67 in FFPE specimens of cervical intraepithelial neoplasia grade I (CIN I). Hematoxylin & Eosin staining (A) with distinct upregulation of ORF1p in the basal two-thirds of the epithelial layer (B), accompanied by a negative p16 staining (C), and basal Ki67 expression pattern (D). Scale bar = 100 μm.

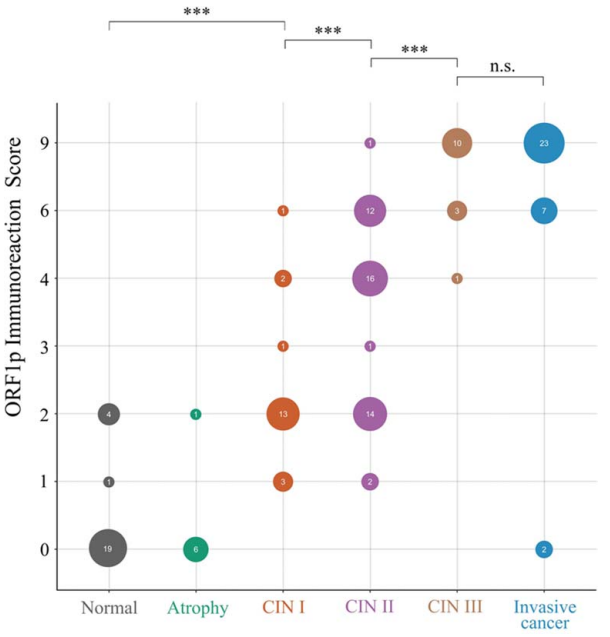

FIGURE 3. Bubble plot to visualize ORF1p immunoreaction scores across different cervical tissue specimens. Bubble sizes and the embedded numbers reflect the number of individuals with a particular immunoreaction score in each tissue sample group. Brackets connect the subjects of contingency analysis (Supplemental Digital Content 1.xls, Supplemental Digital Content 2, http://links.lww.com/IJGP/A177). Three stars indicate $P < 0.001$; n.s, nonsignificant.

and in further cases, the whole cross-section of the epithelium was positive. The extent of ORF1p positivity correlated with CIN grade (Fig. 1). In light of these observations, we have developed a scoring system, which is a modified version of the Immunoreactivity Scale (IRS) by Remmele and Stegner.[42] The final ORF1p immunoreaction score was determined by multiplying the numerical values of staining intensity and epithelial involvement (Table 1 and Materials and Methods section for details). Henceforth, this scoring scheme was used to evaluate the ORF1p immunostaining results of our entire sample panel.

Overall, ORF1p immunopositivity was detected in 98.2% of pre- and neoplastic cervical lesions (score 1–9), while only 20.8% of normal cervical specimens showed weak staining (score 1-2) (Supplemental Digital Content 1.xls, Supplemental Digital Content 2, http://links.lww.com/IJGP/A177). Among normal cervical epithelium samples, only 5 out of 24 displayed mild ORF1p immunoreaction at the basal one or two-thirds of the epithelial thickness (score 1–2), while the remaining 19 samples lacked ORF1p expression (score 0) (Fig. 3). Comparably, only 1 of the 7 atrophic cases exhibited mild ORF1p immunoreaction (score 2) (Fig. 3). In sharp contrast, ORF1p immunopositivity emerged in all CIN I cases, localized predominantly to the lower third of the cervical epithelium with a median value of score 2 (Fig. 3 and Table 2). Of the 20 cases of CIN I, 16 showed a mild immunoreaction (score 1–2), 3 showed a moderate immunoreaction (score 3–4), and 1 showed a strong immunoreaction (score 6) (Fig. 3). Likewise, ORF1p immunopositivity was also present in all CIN II cases, predominantly spanning the lower two-thirds of the cervical epithelium with an increased median score of 4 (Fig. 3 and Table 2). Of the 46 cases of CIN II, 16 showed a mild immunoreaction (score 1–2), 17 showed a moderate immunoreaction (score 3–4) and 13 showed a strong immunoreaction (score 6–9) (Fig. 3). Similarly to the CIN I and CIN II cases, ORF1p immunopositivity was present in all CIN III samples, showing a robust transepithelial appearance with a further increased median score of 9 (Fig. 3 and Table 2). Of the 14 cases of CIN III, only 1 showed a moderate immunoreaction (score 4) while the remaining 13 samples exhibited a strong immunoreaction (score 6–9) (Fig. 3). Samples of invasive cancer cases also exhibited strong ORF1p immunopositivity with a median score of 9 (Fig. 3 and Table 2). 30 of the 32 cancer cases showed strong immunoreaction (score 6–9) (Fig. 3). In the 2 cancer samples where ORF1p expression was absent (Fig. 3), it is conceivable that these tumors shut down LINE-1 expression concurrent with their accelerated growth.[47] The ORF1 immunoreaction scores did not correlate with tumor grade ($P = 0.4989$), and pTNM stage ($P = 0.6039$) (Supplemental Digital Content 1.xls, Supplemental Digital Content 2, http://links.lww.com/IJGP/A177).

Statistical analysis showed stronger immunoreaction for ORF1p in dysplastic lesions (CIN I-III and cancer) versus normal tissue ($P < 0.0001$). A significantly stronger immunoreaction for ORF1p was observed in CIN I versus normal cases ($P < 0.0001$) (Fig. 3. and Supplemental Digital Content 1.xls, Supplemental Digital Content 2, http://links.

TABLE 2. Overview of ORF1p immunoreaction scores in Cervical Tissue Specimens

			ORF1p immunoreaction scores			
	Normal	Atrophy	CIN1	CIN2	CIN3	Invasive cancer
Sample number	24	7	20	46	14	32
Median	0.000	0.000	2.000	4.000	9.000	9.000
Mean	0.375	0.286	2.300	3.870	8.000	7.781
SD	0.770	0.756	1.174	1.809	1.710	2.393

CIN indicates cervical intraepithelial neoplasia.

lww.com/IJGP/A177). In addition, the immunoreaction scores were significantly higher in HSIL than in the CIN I group ($P = 0.001$) (Fig. 3 and Supplemental Digital Content 1.xls, Supplemental Digital Content 2, http://links.lww.com/IJGP/A177). Furthermore, we found statistically significant differences in ORF1p immunoreaction scores between CIN I and CIN II ($P = 0.0075$) and between CIN II and CIN III lesions alike ($P < 0.0001$) (Fig. 3 and Supplemental Digital Content 1.xls, Supplemental Digital Content 2, http://links.lww.com/IJGP/A177). Across our entire sample cohort, no statistically significant correlation was found between age groups (<35, 35–45, 45<) and ORF1p immunopositivity ($P = 0.095$)) (Supplementary Table 1, Supplemental Digital Content 1, http://links.lww.com/IJGP/A176 and Supplemental Digital Content 1.xls, Supplemental Digital Content 2, http://links.lww.com/IJGP/A177). Although, for cases of invasive cancer, we observed slightly elevated median immunoreaction scores in tumor samples from older patients (Supplemental Digital Content 1.xls, Supplemental Digital Content 2, http://links.lww.com/IJGP/A177).

Subsequently, we evaluated the results of the p16 and Ki67 immunostainings for the whole sample panel. To quantify the degree of p16 and Ki67 immunoreactions, we used 3-grade scales, as recommended in references.[13,44,45] As demonstrated by the statistical analyses, both p16 and Ki67 immunoreactions were stronger in the group of all neoplastic and dysplastic lesions (CIN I-III and cancer) versus the control group ($P < 0.001$) (Supplementary Tables 2-3, Supplemental Digital Content 1, http://links.lww.com/IJGP/A176 and Supplemental Digital Content 1.xls, Supplemental Digital Content 2, http://links.lww.com/IJGP/A177). Furthermore, the immunopositivity of these 2 markers graded higher in HSIL than in LSIL ($P < 0.009$) (Supplementary Tables 2-3, Supplemental Digital Content 1, http://links.lww.com/IJGP/A176 and Supplemental Digital Content 1.xls, Supplemental Digital Content 2, http://links.lww.com/IJGP/A177). However, the immunoreaction scores of p16 and Ki67 did not show a significant difference between normal cervical tissue and CIN I ($P = 0.429$ and $P = 0{,}528$) (Supplementary Tables 2-3, Supplemental Digital Content 1, http://links.lww.com/IJGP/A176 and Supplemental Digital Content 1.xls, Supplemental Digital Content 2, http://links.lww.com/IJGP/A177).

Next, we compared the results of ORF1p immunostaining with the results of p16 and Ki67 immunostainings across the whole sample panel. The median immunoreaction scores of the three markers revealed distinct patterns among the various sample groups (Fig. 4). The p16 and Ki67 immunoreaction scores showed a similar profile in the sense that they exhibited a progressive increase from CIN II stage onwards, but remained relatively low in normal, atrophic and CIN I cases (Supplementary Tables 2-3, Supplemental Digital Content 1, http://links.lww.com/IJGP/A176 and Fig. 4). Conversely, ORF1p scores demonstrated a gradual escalation, which was already detected between normal epithelium and CIN I stage samples, and this rising tendency was also observed in the comparison of CIN I and CIN II stages and CIN II and CIN III stages alike (Table 2 and Fig. 4). The immunoreaction scores showed no significant difference between invasive cancer and CIN III sample groups (Table 2 and Fig. 4). Consequently, ORF1p can be regarded as an early marker, serving as useful surrogate for p16 and Ki67, which are markers indicative of later events in the progression of cervical epithelial lesions.

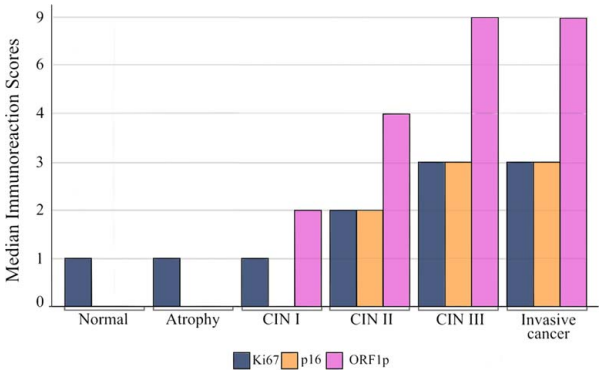

FIGURE 4. The median immunoreaction scores of ORF1p, p16, and Ki67 immunostainings in the entire palette of investigated cervical epithelial lesions. Data shown are derived from the analysis of our whole sample panel (Supplemental Digital Content 1.xls, Supplemental Digital Content 2, http://links.lww.com/IJGP/A177).

DISCUSSION

Our study provides comprehensive insights into the differential expression patterns of ORF1p, p16, and Ki67 markers in cervical dysplasia and cancer. ORF1p emerges as a promising early marker, capable of distinguishing between various CIN grades by offering distinct staining patterns.

In CIN I, we observed subtle ORF1p expression localized predominantly to the lower third of the cervical epithelium. CIN II displayed a moderate increase in ORF1p expression, spanning the lower one and two-thirds of the cervical epithelium. CIN grade 3 and invasive cancer demonstrated robust transepithelial ORF1p immunopositivity, while the normal cervical epithelium consistently lacked ORF1p staining or showed weak positivity (Fig. 1). Compared with the other markers, the greatest practical clinical value of ORF1p immunostaining will probably lie in

its ability to reliably distinguish normal cervical epithelium from CIN I lesions (Figs. 1 and 2). This is evidenced by the mean ORF1p immunoreaction score of 0.37 for normal epithelium and 2.30 for CIN I lesions (Table 2). For CIN I detection, the calculated Positive Predictive Value (PPV) and Negative Predictive Value (NPV) scores (https://www.medcalc.org/calc/diagnostic_test.php) of ORF1p staining were 77% and 100%, respectively, when compared to non-dysplastic samples (Supplemental Digital Content 1.xls, Supplemental Digital Content 2, http://links.lww.com/IJGP/A177). The NPV is particularly favorable because all CIN I cases showed a positive ORF1p immunoreaction. PPV and NPV values were also calculated for other sample groups, which similarly support the usefulness of ORF1p immunostaining (Supplemental Digital Content 1.xls, Supplemental Digital Content 2, http://links.lww.com/IJGP/A177). The utilization of ORF1p as a potential biomarker for identifying CIN lesions bears similarity to the established paradigm wherein ORF1p expression has been exploited as a diagnostic indicator for HPV-independent differentiated vulvar intraepithelial neoplasia (dVIN), the precursor lesion of vulval squamous cell carcinoma.[43]

Besides, ORF1p immunostaining provides a clearer and more reliable distinction between higher CIN grades than the conventionally used p16 and Ki67 markers (Figs. 3 and 4). While p16 immunostaining has proven valuable in identifying high-risk HPV-associated lesions, its interpretation can sometimes be subjective and prone to interobserver variability.[48] This is also illustrated by the fact that the authors of Clark et al[19] found that 19% of the examined HSIL samples had to be reclassified as LSIL because positive p16 immunostainings were overvalued at the time of the original diagnosis. In combination with ORF1p immunostaining, the concomitant presence of a mild ORF1p immunoreaction characteristic of LSIL may help the investigator to avoid such an error due to the presence of positive p16 immunostaining. Our current study suggests that in HSIL cases, the moderate or strong ORF1p immunoreaction is much more typical.

Our present results strongly support that the retrotransposition activity of endogenous LINE-1 elements may be an important driver of cervical tumorigenesis from its very beginning, as evidenced by the generalized ORF1p immunopositivity already in CIN I stage (Figs. 1 and 3). Over the last decade, it has become clear that the appearance of LINE-1 proteins and the concomitant LINE-1 retrotransposition activity in somatic cells is not only a simple marker of tumorigenesis but is also a driver of that process, that is, a pathogenic factor. It was shown that LINE-1 activity is present in more than 50% of cancers[29,35] and is now recognized as a major hallmark of cancer accompanied by genomic instability and genetic heterogeneity.[49] In line with this, several cancer driver mutations were identified in different tumor types that were caused by novel somatic LINE-1 integration events.[29,30,41,50–53] Moreover, the targeted inhibition of the activity of LINE-1 elements showed a clinical benefit in a phase 2 clinical trial in the treatment of metastatic colorectal carcinoma.[54] Notably, in the cervical epithelium, there is a potential interplay between LINE-1 retrotransposition and HPV infection. HPV infection may lead to alterations in the epigenetic landscape of infected cells.[55] In turn, HPV-induced epigenetic modifications may contribute to the activation of LINE-1 retrotransposition.[32] This is underlined by the fact that LINE-1 elements were found to be differentially expressed among cervical cancer samples infected with different HPV serotypes.[56]

Invasive cervical cancer cases, similar to CIN III lesions, exhibited a strong ORF1p immunoreaction (median score 9) but with some degree of intra-tumor heterogeneity (Fig. 1, Fig. 3, and Table 2). Remarkably, the authors of Ardeljan et al proposed that LINE-1 retrotransposition may conflict with DNA replication in exceptionally fast-growing tumors or tumor subclones.[47] In other words, exceptionally fast-dividing cancer cells may benefit from the shutdown of the LINE-1 retrotransposition activity that promoted earlier stages of their transformation. This may explain both the intra-tumor heterogeneity of ORF1p immunoreaction and the presence of 2 cancer samples where ORF1p expression was absent (Figs. 1 and 3).

Rarely, we also saw sporadic and weak ORF1p immunopositivity in normal cervical epithelium (data not shown). This is validated by the recent demonstration of McKerrow et al that LINE-1 mRNA expression is sometimes detectable in normal epithelial cells across various tissue types.[28] This suggests that LINE-1 expression is not completely sealed in epithelial cells and leaky expression may occur. Epigenetic downregulation of LINE-1 transcription is the first line of defense against mutagenic somatic retrotransposition. At the organism level, LINE-1 transcription is suppressed by epigenetic mechanisms already in the germline and early embryo due to the action of the piRNA system.[57,58] Heterochromatin established on LINE-1 loci in early developmental stages is then passed on to somatic tissues during individual development by cell divisions to maintain the transcriptionally repressed state of LINE-1 elements. However, sporadically in epithelial cells, leaky transcription of some LINE-1 copies can occur potentially in response to stimuli that may interfere with these epigenetic mechanisms. Presumably, this phenomenon is what we and McKerrow and colleagues have observed in epithelial cells. For this reason, we classified ORF1p immunostaining as negative when only less than 10% weakly reactive epithelial cells were present (see Materials and Methods). Future research is needed to show whether such weak sporadic expression observed in normal epithelium leads to new retrotransposition events, as there are many posttranscriptional and post-translational mechanisms for LINE-1 regulation[31,32] that may yet be able to prevent this. Although, it is clear that the stronger LINE-1 expression observed in tumors leads to retrotransposition events.[29]

The potential utility of ORF1p as a solid tumor biomarker is demonstrated by the fact that its detection for diagnostic purposes is currently even being considered from body fluids. Recently, Taylor et al have demonstrated that ORF1p detection from plasma may serve as a

multi-cancer biomarker with potential utility for disease detection and monitoring.[59] Another recent study by Sato et al have revealed the detectability of ORF1p in ascites and plasma samples from patients with high-grade serous ovarian cancer, also underscoring the potential of ORF1p as a promising candidate serum biomarker.[38]

While our study provides promising insights into ORF1p as a novel biomarker for CIN diagnosis, it is important to acknowledge certain limitations. The relatively small sample size, especially in specific subgroups, might impact the statistical power and robustness of the findings. In addition, the utilization of Tissue Microarray technology, despite our efforts to ensure core representativeness, introduces a limitation in capturing the full heterogeneity within complex lesions. It is noteworthy that ORF1p staining exhibits a more uniform distribution within cervical lesions compared with the often segmental and patchy patterns seen in p16 staining. We established that 2 mm tissue cores from the selected target lesions were most appropriate for constructing our TMAs, which is in line with previous studies.[60–63] Two-millimeter cores are large enough to minimize problems of representativity but small enough to allow all samples to be treated with identical staining protocol during analysis. While our findings are promising, further large-scale studies are essential to fully validate the diagnostic and prognostic utility of ORF1p across diverse demographics and clinical settings.

Our findings support that the use of ORF1p alongside Ki67 and p16 markers enhances the diagnostic capability and potential reproducibility in distinguishing dysplasia from normal epithelium and differentiation among the CIN grades. ORF1p appears as a valuable candidate biomarker that offers a more reliable diagnosis for improved CIN patient management.

ACKNOWLEDGMENTS

The authors thank Ivett Csikós and Sarolta Bankó for technical assistance and administrative support. They also thank Tibor Kun, Beáta Balogh, and Krisztián Daru for their excellent immunohistochemistry work and Mihály Dezső for photography.

REFERENCES

1. *WHO guideline for screening and treatment of cervical pre-cancer lesions for cervical cancer prevention: use of mRNA tests for human papillomavirus (HPV)*. 2021, World Health Organization.
2. *Cervical cancer statistics*. World Cancer Research Fund (WCRF) International; 2020.
3. Moreno-Acosta P, Romero-Rojas A, Vial N, et al. Persistent high-Risk HPV infection and molecular changes related to the development of cervical cancer. *Case Rep Obstet Gynecol* 2020;2020:6806857.
4. Ambros RA, Kurman RJ. Current concepts in the relationship of human papillomavirus infection to the pathogenesis and classification of precancerous squamous lesions of the uterine cervix. *Semin Diagn Pathol* 1990;7:158–172.
5. Kalliala I, Athanasiou A, Veroniki AA, et al. Incidence and mortality from cervical cancer and other malignancies after treatment of cervical intraepithelial neoplasia: A systematic review and meta-analysis of the literature. *Ann Oncol* 2020;31:213–227.
6. Chan CK, Aimagambetova G, Ukybassova T, et al. Human papillomavirus infection and cervical cancer: epidemiology, screening, and vaccination-review of current perspectives. *J Oncol* 2019; 2019:3257939.
7. Stoler MH, Schiffman M. Atypical squamous cells of undetermined significance-low-grade squamous intraepithelial lesion triage study. *Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ASCUS-LSIL Triage Study. JAMA* 2001;285:1500–1505.
8. Creagh T, Bridger JE, Kupek E, et al. Pathologist variation in reporting cervical borderline epithelial abnormalities and cervical intraepithelial neoplasia. *J Clin Pathol* 1995;48:59–60.
9. Carreon JD, Sherman ME, Guillén D, et al. CIN2 is a much less reproducible and less valid diagnosis than CIN3: Results from a histological review of population-based cervical samples. *Int J Gynecol Pathol* 2007;26:441–446.
10. Gage JC, Schiffman M, Hunt WC, et al. Cervical histopathology variability among laboratories: A population-based statewide investigation. *Am J Clin Pathol* 2013;139:330–335.
11. Silva DC, Gonçalves AK, Cobucci RN, et al. Immunohistochemical expression of p16, Ki-67 and p53 in cervical lesions - A systematic review. *Pathol Res Pract* 2017;213:723–729.
12. Bullwinkel J, Baron-Lühr B, Lüdemann A, et al. Ki-67 protein is associated with ribosomal RNA transcription in quiescent and proliferating cells. *J Cell Physiol* 2006;206:624–635.
13. Nam EJ, Kim JW, Hong JW, et al. Expression of the p16 and Ki-67 in relation to the grade of cervical intraepithelial neoplasia and high-risk human papillomavirus infection. *J Gynecol Oncol* 2008;19:162–168.
14. Ordi J, Garcia S, del Pino M, et al. p16 INK4a immunostaining identifies occult CIN lesions in HPV-positive women. *Int J Gynecol Pathol* 2009;28:90–97.
15. O'Neill CJ, McCluggage WG. p16 expression in the female genital tract and its value in diagnosis. *Adv Anat Pathol* 2006;13:8–15.
16. Mandal R, Ghosh I, Banerjee D, et al. Correlation between p16/Ki-67 expression and the grade of cervical intraepithelial neoplasias. *Int J Gynecol Pathol* 2020;39:384–390.
17. Huang EC, Tomic MM, Hanamornroongruang S, et al. p16ink4 and cytokeratin 7 immunostaining in predicting HSIL outcome for low-grade squamous intraepithelial lesions: A case series, literature review and commentary. *Mod Pathol* 2016;29:1501–1510.
18. Jackson JA, Kapur U, Ersahin C. Utility of p16, Ki-67, and HPV test in diagnosis of cervical intraepithelial neoplasia and atrophy in women older than 50 years with 3- to 7-year follow-up. *Int J Surg Pathol* 2012;20:146–153.
19. Clark JL, Lu D, Kalir T, et al. Overdiagnosis of HSIL on cervical biopsy: Errors in p16 immunohistochemistry implementation. *Hum Pathol* 2016;55:51–56.
20. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature* 2001;409:860–921.
21. Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. *Nat Rev Genet* 2009;10:691–703.
22. Eickbush T. Exon shuffling in retrospect. *Science* 1999;283:1465; 1467.
23. Moran JV, DeBerardinis RJ, Kazazian HH Jr. Exon shuffling by L1 retrotransposition. *Science* 1999;283:1530–1534.
24. Hancks DC, Kazazian HH Jr. Active human retrotransposons: Variation and disease. *Curr Opin Genet Dev* 2012;22:191–203.
25. Ostertag EM, DeBerardinis RJ, Goodier JL, et al. A mouse model of human L1 retrotransposition. *Nat Genet* 2002;32:655–660.
26. Kano H, Godoy I, Courtney C, et al. L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. *Genes Dev* 2009;23:1303–1312.
27. Xing J, Zhang Y, Han K, et al. Mobile elements create structural variation: Analysis of a complete human genome. *Genome Res* 2009; 19:1516–1526.
28. McKerrow W, Kagermazova L, Doudican N, et al. LINE-1 retrotransposon expression in cancerous, epithelial and neuronal cells revealed by 5' single-cell RNA-Seq. *Nucleic Acids Res* 2023;51: 2033–2045.
29. Rodriguez-Martin B, Alvarez EG, Baez-Ortega A, et al. Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by LINE-1 retrotransposition. *Nat Genet* 2020;52:306–319.
30. Cajuso T, Sulo P, Tanskanen T, et al. Retrotransposon insertions can initiate colorectal cancer and are associated with poor survival. *Nat Commun* 2019;10:4022.

31. Pizarro JG, Cristofari G. Post-transcriptional control of LINE-1 retrotransposition by cellular host factors in somatic cells. *Front Cell Dev Biol* 2016;4:14.
32. Protasova MS, Andreeva TV, Rogaev EI. Factors regulating the activity of LINE1 retrotransposons. *Genes (Basel)* 2021;12:1562.
33. Martin SL, Cruceanu M, Branciforte D, et al. LINE-1 retrotransposition requires the nucleic acid chaperone activity of the ORF1 protein. *J Mol Biol* 2005;348:549–561.
34. Feng Q, Moran JV, Kazazian HH Jr, et al. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* 1996;87:905–916.
35. Rodic N, Sharma R, Sharma R, et al. Long interspersed element-1 protein expression is a hallmark of many human cancers. *Am J Pathol* 2014;184:1280–1286.
36. Whongsiri P, Pimratana C, Wijitsettakul U, et al. LINE-1 ORF1 protein is up-regulated by reactive oxygen species and associated with bladder urothelial carcinoma progression. *Cancer Genomics Proteomics* 2018;15:143–151.
37. Chen L, Dahlstrom JE, Chandra A, et al. Prognostic value of LINE-1 retrotransposon expression and its subcellular localization in breast cancer. *Breast Cancer Res Treat* 2012;136:129–142.
38. Sato S, Gillette M, de Santiago PR, et al. LINE-1 ORF1p as a candidate biomarker in high grade serous ovarian carcinoma. *Sci Rep* 2023;13:1537.
39. Xia Z, Cochrane DR, Tessier-Cloutier B, et al. Expression of L1 retrotransposon open reading frame protein 1 in gynecologic cancers. *Hum Pathol* 2019;92:39–47.
40. Rodic N. LINE-1 retrotransposons as neoplastic biomarkers, *Human Retrotransposons in Health and Disease*. Cristofari G. Editor. 2017 Springer International Publishing, Cham. 275–295.
41. Doucet-O'Hare TT, Rodić N, Sharma R, et al. LINE-1 expression and retrotransposition in Barrett's esophagus and esophageal carcinoma. *Proc Natl Acad Sci USA* 2015;112:E4894–E4900.
42. Remmele W, Stegner HE. [Recommendation for uniform definition of an immunoreactive score (IRS) for immunohistochemical estrogen receptor detection (ER-ICA) in breast cancer tissue]. *Pathologe* 1987; 8:138–140.
43. Hofstetter G, Mildner M, Tschandl P, et al. ORF1p is a potential novel diagnostic marker for differentiated vulvar intraepithelial neoplasia. *Int J Gynecol Pathol* 2023;42:201–206.
44. Klaes R, Friedrich T, Spitkovsky D, et al. Overexpression of p16 (INK4A) as a specific marker for dysplastic and neoplastic epithelial cells of the cervix uteri. *Int J Cancer* 2001;92:276–284.
45. Alshenawy HA. Evaluation of p16, human papillomavirus capsid protein L1 and Ki-67 in cervical intraepithelial lesions: Potential utility in diagnosis and prognosis. *Pathol Res Pract* 2014;210:916–921.
46. R Core Team (2023). *R: A Language and Environment for Statistical Computing* 2023R Foundation for Statistical Computing, Vienna, Austria.
47. Ardeljan D, Steranka JP, Liu C, et al. Cell fitness screens reveal a conflict between LINE-1 retrotransposition and DNA replication. *Nat Struct Mol Biol* 2020;27:168–178.
48. van Bogaert LJ. P16INK4a immunocytochemistry/immunohistochemistry: Need for scoring uniformization to be clinically useful in gynecological pathology. *Ann Diagn Pathol* 2012;16:422–426.
49. Zhang X, Zhang R, Yu J. New understanding of the relevant role of LINE-1 retrotransposition in human disease and immune modulation. *Front Cell Dev Biol* 2020;8:657.
50. Shukla R, Upton KR, Muñoz-Lopez M, et al. Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma. *Cell* 2013;153:101–111.
51. Ewing AD, Gacita A, Wood LD, et al. Widespread somatic L1 retrotransposition occurs early during gastrointestinal cancer evolution. *Genome Res* 2015;25:1536–1545.
52. Rodic N, Steranka JP, Makohon-Moore A, et al. Retrotransposon insertions in the clonal evolution of pancreatic ductal adenocarcinoma. *Nat Med* 2015;21:1060–1064.
53. Miki Y, Nishisho I, Horii A, et al. Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. *Cancer Res* 1992;52:643–645.
54. Rajurkar M, Parikh AR, Solovyov A, et al. Reverse transcriptase inhibition disrupts repeat element life cycle in colorectal cancer. *Cancer Discov* 2022;12:1462–1481.
55. Mac M, Moody CA. Epigenetic regulation of the human papillomavirus life cycle. *Pathogens* 2020;9:483.
56. Curty G, Menezes AN, Brant AC, et al. Expression of retroelements in cervical cancer and their interplay with HPV infection and host gene expression. *Cancers (Basel)* 2021;13:3513.
57. Pezic D, Manakov SA, Sachidanandam R, et al. piRNA pathway targets active LINE1 elements to establish the repressive H3K9me3 mark in germ cells. *Genes Dev* 2014;28:1410–1428.
58. Marchetto MCN, Narvaiza I, Denli AM, et al. Differential L1 regulation in pluripotent stem cells of humans and apes. *Nature* 2013; 503:525–529.
59. Taylor MS, Connie W, Fridy PC, et al, *Ultrasensitive detection of circulating LINE-1 ORF1p as a specific multi-cancer biomarker.* bioRxiv, 2023. doi:10.1101/2023.01.25.525462
60. Lesnikova I, Lidang M, Hamilton-Dutoit S, et al. p16 as a diagnostic marker of cervical neoplasia: A tissue microarray study of 796 archival specimens. *Diagn Pathol* 2009;4:22.
61. He G, Chen L, Ye Y, et al. Piwil2 expressed in various stages of cervical neoplasia is a potential complementary marker for p16. *Am J Transl Res* 2010;2:156–169.
62. Popiel A, Piotrowska A, Sputa-Grzegrzolka P, et al. Preliminary study on the expression of testin, p16 and Ki-67 in the Cervical Intraepithelial Neoplasia. *Biomedicines* 2021;9:1010.
63. Tawfik El-Mansi M, Williams AR. Validation of tissue microarray technology using cervical adenocarcinoma and its precursors as a model system. *Int J Gynecol Cancer* 2006;16:1225–1233.

II.

BMC Biology

METHODOLOGY ARTICLE

Open Access

A versatile transposon-based technology to generate loss- and gain-of-function phenotypes in the mouse liver

Anna Georgina Kopasz[1†], Dávid Zsolt Pusztai[1,2†], Réka Karkas[1,3], Liza Hudoba[1], Khaldoon Sadiq Ahmed Abdullah[1,3], Gergely Imre[1,2], Gabriella Pankotai-Bodó[4], Ede Migh[5], Andrea Nagy[1], András Kriston[5], Péter Germán[1], Andrea Bakné Drubi[1,2], Anna Molnár[1], Ildikó Fekete[1], Virág Éva Dani[1], Imre Ocsovszki[6], László Géza Puskás[1], Péter Horváth[5,7], Farkas Sükösd[4] and Lajos Mátés[1*]

Abstract

Background Understanding the contribution of gene function in distinct organ systems to the pathogenesis of human diseases in biomedical research requires modifying gene expression through the generation of gain- and loss-of-function phenotypes in model organisms, for instance, the mouse. However, methods to modify both germline and somatic genomes have important limitations that prevent easy, strong, and stable expression of transgenes. For instance, while the liver is remarkably easy to target, nucleic acids introduced to modify the genome of hepatocytes are rapidly lost, or the transgene expression they mediate becomes inhibited due to the action of effector pathways for the elimination of exogenous DNA. Novel methods are required to overcome these challenges, and here we develop a somatic gene delivery technology enabling long-lasting high-level transgene expression in the entire hepatocyte population of mice.

Results We exploit the fumarylacetoacetate hydrolase (Fah) gene correction-induced regeneration in Fah-deficient livers, to demonstrate that such approach stabilizes luciferase expression more than 5000-fold above the level detected in WT animals, following plasmid DNA introduction complemented by transposon-mediated chromosomal gene transfer. Building on this advancement, we created a versatile technology platform for performing gene function analysis in vivo in the mouse liver. Our technology allows the tag-free expression of proteins of interest and silencing of any arbitrary gene in the mouse genome. This was achieved by applying the HADHA/B endogenous bidirectional promoter capable of driving well-balanced bidirectional expression and by optimizing in vivo intronic artificial microRNA-based gene silencing. We demonstrated the particular usefulness of the technology in cancer research by creating a p53-silenced and hRas G12V-overexpressing tumor model.

Conclusions We developed a versatile technology platform for in vivo somatic genome editing in the mouse liver, which meets multiple requirements for long-lasting high-level transgene expression. We believe that this technology will contribute to the development of a more accurate new generation of tools for gene function analysis in mice.

Keywords *Fah* KO mouse, *Sleeping Beauty*, Somatic transgenesis, In vivo gene silencing, Tumor model

†Anna Georgina Kopasz and Dávid Zsolt Pusztai are joint first authors.

*Correspondence:
Lajos Mátés
mates.lajos@brc.hu
Full list of author information is available at the end of the article

Background

Genetic manipulations to modify gene expression in a cell or an organism have been applied in numerous fields including research, medicine, industrial

biotechnology, and agriculture. In research, they are used to study gene function through the generation of loss-of-function and gain-of-function phenotypes, to characterize gene expression patterns through the introduction of reporter genes, and to visualize intracellular trafficking of macromolecules through mRNA and protein tagging. The laboratory mouse is currently the predominant mammalian species in biomedical research, most commonly used as an experimental model system for investigating the pathogenesis of human diseases and for developing new therapies. However, genetic manipulations in the mouse germline are generally laborious and time consuming and frequently result in unstable transgene expression or unreliable spatiotemporal expression pattern [1, 2]. Alternatively, if genetic modifications target the soma of mice, the workflows are faster. However, these workflows are becoming dominated by viral gene delivery, and in turn, they are hampered by viral cargo limitations, and despite natural viral infectivity, they often result in unstable gene expression due to host immune response against viral proteins [3, 4]. To replace viruses, lipid nanoparticles are used extensively as synthetic non-viral delivery vehicles. They are less prone to trigger host immune response than viral vectors due to the absence of immunogenic viral proteins and do not exhibit strict cargo limitations. However, lipid nanoparticle-based somatic gene delivery is generally not as effective as the viral one. Different organs can be targeted with varying degrees of efficiency, but in the case of the liver, the procedure is particularly effective [5]. The liver can also be efficiently targeted with naked plasmid DNA using a simple in vivo transfection procedure called hydrodynamic injection [6].

However, transgene expression rapidly declines in the liver following plasmid DNA delivery [7]. To date, a number of receptors have been reported to recognize cytosolic exogenous DNA, such as Toll-like receptor 9 (TLR9), cyclic GMP-AMP synthase (cGAS), and absent in melanoma 2 (AIM2) inflammasomes [8, 9]. These receptors trigger effector pathways that contribute to the elimination and transcriptional repression of plasmid DNA. To improve the outcome of plasmid DNA delivery, the system can be supplemented with non-viral transposon-based chromosomal gene transfer. Indeed, together with *Sleeping Beauty* (SB) transposon-mediated chromosomal transfer of the transgene, the long-term parameters of gene expression slightly improve, but a few weeks after injection, the decrease in transgene expression remains enormous [10]. This suggests that the effect of cytosolic DNA sensors can only be partially avoided, even if the technology is complemented by chromosomal gene transfer. Problems persist and transgene expression is not stably maintained, at least not to a sufficient extent.

Our aim was to change this situation by improving the current biotechnology toolkit and harnessing the SB transposon together with an efficient somatic transgenesis system, taking the efficiency, versatility, and stability of liver-specific gene delivery in mice to the next level. We planned to create a technology that would simultaneously allow the expression of native or mutant isoforms of proteins and efficient silencing of any arbitrary target gene in the mouse genome, in a stable manner and a high number of cells. We took the advantage of hydrodynamic injection for efficient in vivo transfection of hepatocytes and hyperactive SB [11] transposon-mediated chromosomal gene transfer for stable transgene delivery. For achieving high affected cell number and high stability of transgene expression, we harnessed the known selection pressure exerted in fumarylacetoacetate hydrolase (*Fah*) KO livers for Fah-expressing hepatocytes [12]. Finally, for more versatile gene expression modifications, we applied a new bidirectional promoter that had not previously been part of the biotechnology toolkit and optimized intron-derived microRNA (miR)-based gene silencing in vivo in the mouse liver.

The technology platform reported here is well suited to study gene function in hepatocytes via the generation of gain-of-function and loss-of-function phenotypes as it allows the exchange of virtually the entire hepatocyte population for transgenic cells. Here we demonstrate its particular usefulness in cancer research. A current high priority in cancer research is to functionally validate candidate genetic alterations that are relevant for cancer progression as it is no longer possible to clearly identify them using bioinformatics methods based on mutation frequency analysis alone, due to potential candidates being mutated at lower frequencies in cancer samples. Thus, there is a growing demand for in vivo experimental systems where the cancer driver role of mutations could be confirmed.

Results

Simple and highly efficient gene delivery system for directed gene expression modifications in the mouse liver

Keeping in mind the general simplicity of the procedure, we based our system on the hydrodynamic injection of plasmid DNA constructs. Hydrodynamic plasmid delivery primarily targets hepatocytes by the enhancement of their membrane permeability [6, 13]. However, following in vivo transfection, chromosomal integration is also required for long-term stable transgene expression. For very efficient SB transposon-based chromosomal transgene delivery, we used a hyperactive transposase

helper SB100 [11] and a more active transposon, the so-called T2 Inverted Terminal Repeat (ITR) structure variant [14]. In those hepatocytes where hydrodynamic transfection is successful, the hyperactive transposase helper enzyme is likely to catalyze the "cut and paste" transposition reaction presumably leading to an integration into the host chromosomes [15]. However, literature data suggest that the negative effect of exogenous DNA sensors on transgene expression cannot be efficiently avoided, even if the technology is complemented by chromosomal gene transfer [10]. We hypothesized that a high level of sustained transgene expression could be achieved by exploiting the well-known high regenerative potential of adult mouse hepatocytes. Following an extended hepatectomy, which involves removing nearly 90% of liver tissue, the mouse liver can regenerate within a short time regaining its normal size [16]. Such a high regenerative potential can be harnessed to replace virtually all $Fah^{-/-}$ hepatocytes for $Fah^{+/+}$ ones in a *Fah*-deficient liver by multi-nodular repopulation due to the selective growth advantage of wild-type (WT) cells [12]. Therefore, we supplemented our transposon construct with a *Fah* coding sequence (CDS) and used a *Fah* mutant mouse strain (C57BL/6N-$Fah^{tm1(NCOM)Mfgc/Biat}$) for the hydrodynamic injections (Fig. 1a).

For directed gene expression modifications, we constructed an SB transposon-based cloning platform (Fig. 1a) built on the co-expression of two linked transcripts allowing transgene expression to be bound to the expression of the *Fah* selection marker. Transcript A codes for the selection and visualization marker proteins connected by a "self-cleaving" T2A peptide for bicistronic expression [17], whereas transcript B encodes the protein to be tested without the need for tagging. To this end, we complemented the current biotechnology toolbox with the human *HADHA/B* promoter capable of driving well-balanced bidirectional expression. The compact (390 bp) *HADHA/B* bidirectional promoter drives the expression of the human hydroxyacyl-CoA dehydrogenase trifunctional multienzyme complex alpha (HADHA) and beta (HADHB) subunits [18].

To test the functionality of the design, an EGFP-encoding transcript B driven by the *HADHB* side of the bidirectional promoter was created (Fig. 1a). First, this EGFP-expressing transposon and the SB100 transposase helper plasmid were co-delivered hydrodynamically into the liver of $Fah^{-/-}$ mice, then the curative drug nitisinone (NTBC) [12] that had been given continuously up to that point was withdrawn. Immunohistochemical (IHC) investigations (Fig. 1b) demonstrated that due to intensive multi-nodular repopulation after 3 months, virtually all hepatocytes were Fah- and EGFP-positive in the treated livers. Next, using a quantitative reverse transcription PCR (RT-qPCR) method, we identified the amount of transcripts A and B produced by the *HADHA* and *HADHB* sides of the promoter, respectively. The amounts of mCherry-T2A-Fah and EGFP mRNAs were normalized to that of the ribosomal protein L27 (*Rpl27*). Our RT-qPCR data demonstrate that the transgene had an order of magnitude stronger expression than the ribosomal protein Rpl27 in the liver (Fig. 1c), which is fully consistent with a strong but physiological gene expression level. It should be noted that the expression levels of transcripts A and B were virtually identical (1/1.18).

To better explore the impact of multi-nodular repopulation on long-term transgene expression in *Fah*-deficient livers, the same construct, in which transcript B was coding for the firefly luciferase (Luc) marker protein (Fig. 1a), was injected into both wild-type (WT) and *Fah* KO mice in the presence of the SB100 transposase helper. Transgene expression was monitored over time by detecting the bioluminescence in living animals using the IVIS Lumina III imaging system (PerkinElmer). After a slightly higher Luc expression detected in WT animals during the first week after hydrodynamic injection, this transgene expression trend was reversed from day 28 onwards. In 3 months, by day 84, the average bioluminescence intensity in WT animals decreased to 387-fold of the initial level, while in *Fah* KO animals it reached 61-fold of the initial level (Fig. 1d, e).

Design and in vivo application of amiR elements

Our SB transposon-based cloning platform also provides an opportunity for simultaneous silencing of endogenous genes by the expression of artificial microRNA (amiR) elements. In order to avoid interference with the translation of marker proteins, amiR structures were expressed from an intron inserted into transcript A. To accommodate intronic amiR structures, the modified 943-bp-long first intron of the human eukaryotic translation elongation factor 1 alpha 1 (*EEF1A1*) gene was incorporated in the mCherry CDS (Fig. 1a). To verify the in vivo applicability of intronic amiR elements in our platform, we first incorporated an amiR element silencing EGFP (amiR-EGFP) into the intron of the mCherry CDS. In this arrangement, amiR-EGFP, which is processed from transcript A, silences EGFP encoded by transcript B. The guide sequence of amiR-EGFP targeting the EGFP CDS was designed by Beisel et al. [19]. For the creation of the amiR-EGFP, "miR-E," an optimized human miR-30a-based miR backbone, was applied [20], and the EGFP guide sequence was inserted into the miR-E backbone structure (Additional file 1: Fig. S1). This amiR-EGFP- and EGFP-expressing transposon and the SB100 transposase helper plasmid were co-delivered into the liver of $Fah^{-/-}$ mice, then NTBC was withdrawn. After

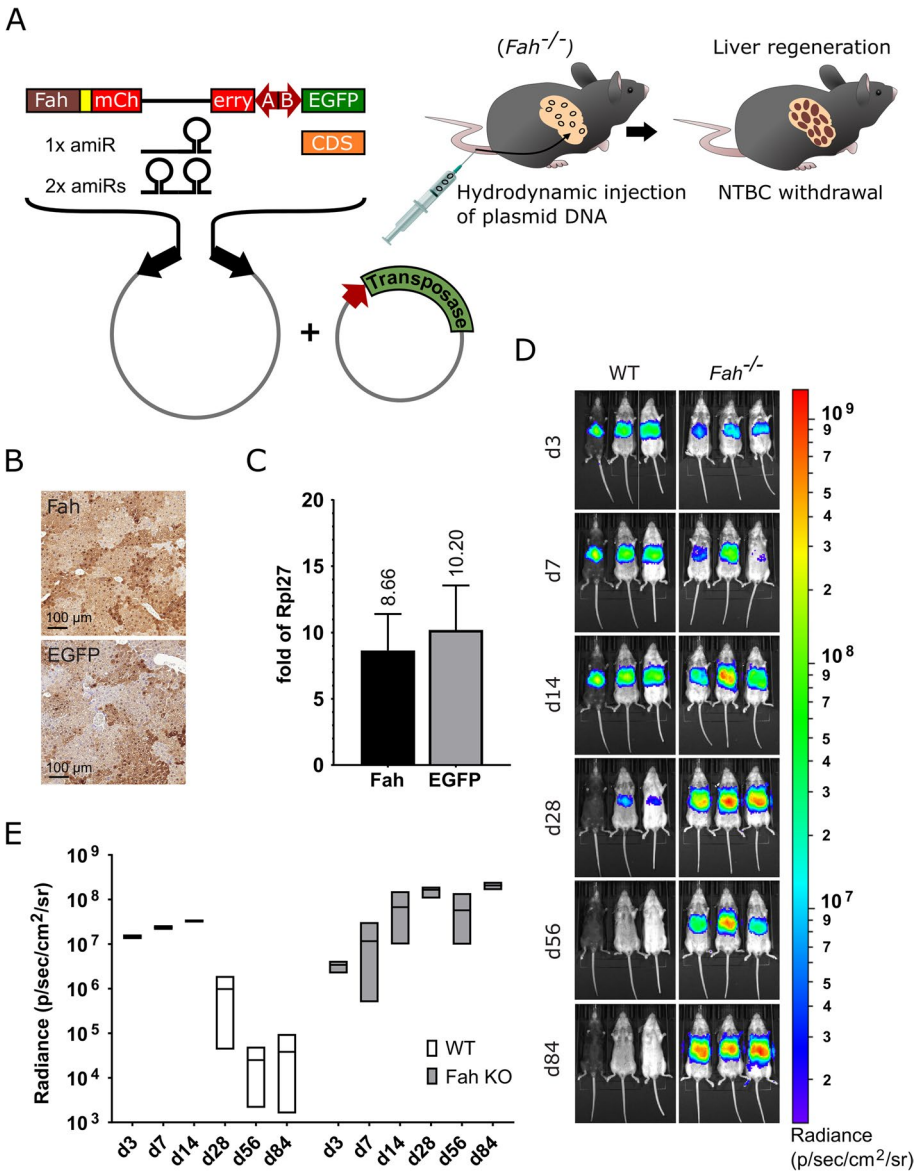

Fig. 1 In vivo transposon-based gene delivery into the liver of *Fah*$^{-/-}$ and WT mice. **a** Schematic representation of the *Sleeping Beauty* (SB) transposon-based cloning platform and animal treatments. Black arrows, SB transposon inverted terminal repeats; red arrows, promoters. **b** Fah and EGFP immunostainings of liver sections from *Fah*$^{-/-}$ mice 3 months after NTBC withdrawal. **c** Monitoring the amount of transcripts A and B following in vivo gene delivery. Liver RNA samples were collected from *Fah*$^{-/-}$ mice at 3 months post-treatment. Samples were tested using Fah- and EGFP-specific RT-qPCR assays. Results were normalized to measurements of the ribosomal protein L27 (*Rpl27*) transcript as input control and data were presented as the mean ± standard deviation (SD) ($n = 3$) (see Additional file 2 for individual data values and statistics). **d** Live bioluminescence imaging of *Fah*$^{-/-}$ and WT mice following in vivo gene delivery. Bioluminescence signals were obtained using an IVIS Lumina III imaging system at 3, 7, 14, 28, 56, and 84 days post-treatment. **e** Kinetics of bioluminescence changes during the first 3 months after gene delivery. For each experimental animal, the average radiance (photons/second/cm^2/steradian (sr) [p/s/cm^2/sr]) of circular regions of the same size covering the liver area was used for plotting. The numerical values were presented as box diagram from lowest to highest values with line at mean ($n = 3$) (see Additional file 2 for individual data values and statistics)

5 months, following complete multi-nodular repopulation, the animals were sacrificed. Macrovisualization of mCherry and EGFP autofluorescence in the liver showed that silencing of EGFP expression was very effective compared to control animals without amiR structures (Fig. 2a). The results of our EGFP Western blot assays also confirmed effective knockdown at the protein level (Additional file 1: Fig. S2). Next, RT-qPCR measurements

were used to determine the amount of A and B transcripts. According to these measurements, transcript B encoding EGFP was silenced to 4% residual gene expression in livers expressing amiR-EGFP elements (Fig. 2b).

After successful in vivo silencing of EGFP, we selected an endogenous target gene, *Tp53*, for amiR-mediated silencing. For amiR-mP53 structures — similarly to amiR-EGFP — guide sequences targeting endogenous p53 mRNA were inserted into the "miR-E" backbone (Additional file 1: Fig. S1). The amiR-mP53 guide sequences were designed following the guide design rules of Dow et al. [21] summarized in Additional file 1: Fig. S1.

To test the efficacy of amiR-mP53 elements designed to silence mouse *Tp53*, we performed studies in a mammalian tissue culture system. To this end, we inserted the first intron of the EEF1A1 gene, modified to accommodate amiR elements, into the sequence encoding the Neomycin selection marker protein and then incorporated the three amiR-mP53 elements one by one into this intron. The resulting Neomycin expression units were then inserted between the SB transposon ITRs (Additional file 1: Fig. S3). Next, the transposon series carrying the different amiR-mP53 elements and an amiR-free control transposon were co-transfected with the SB100 transposase helper plasmid into NIH3T3 cells. After G418 selection, RT-qPCR measurements were performed to determine the level of endogenous p53 mRNA in the transfected cells. Our measurements showed that all three amiR-mP53 elements were sufficiently effective. However, amiR-mP53/1 and amiR-mP53/2 elements performed better, producing 20.34 and 7.17% remaining gene expression, respectively, and were used for further in vivo studies (Fig. 2c). Subsequently, amiR-mP53/1 alone or amiR-mP53/1 and amiR-mP53/2 elements together were incorporated into the intron of the mCherry CDS in our cloning platform designed for in vivo studies (Fig. 1a). These amiR-mP53-expressing transposons and the SB100 transposase helper plasmid were co-delivered hydrodynamically into the liver of $Fah^{-/-}$ mice, then NTBC was withdrawn. After 5 months, following complete multi-nodular repopulation, the animals were sacrificed. Macrovisualization of the mCherry and EGFP autofluorescence showed that in these organs the expression of the mCherry and EGFP marker proteins was balanced and comparable to the control (Fig. 2a). Our RT-qPCR measurements confirmed that transcript A and B levels were not significantly altered compared to the control (Fig. 2b). However, we found 53% remaining *Tp53* expression in organs expressing amiR-mP53/1, while organs co-expressing amiR-mP53/1 and amiR-mP53/2 showed 32% remaining *Tp53* expression (Fig. 2b). This is in line with expectations, as hepatocytes make up about half of liver tissue [22], the other cell types are not affected by *Tp53* silencing. In contrast, EGFP is expressed exclusively in hepatocytes and therefore all EGFP-expressing cells are affected by gene silencing.

With successful gene silencing, the liver regeneration of animals expressing amiR was similar to that of controls without amiR, and no signs of non-specific toxicity associated with amiR expression were observed. The slight decrease in transgene copy number observed in livers expressing one or two amiR-mP53 elements (Fig. 2d) may explain the slight, non-significant decrease in transcript A and B levels in these organs as compared to controls (Fig. 2b).

Hepatocellular carcinoma modeling using predefined combinations of drivers

To demonstrate the utility of our technology in cancer research and the ability to co-express mutant or native proteins with amiR elements, in our cloning platform, we created an SB transposon armed to silence the endogenous *Tp53* and overexpress an oncogenic hRas variant. To achieve this, amiR-mP53/1 was introduced into the intron of the mCherry CDS in transcript A and $hRas^{G12V}$ — a constitutively active form of hRas [23] — was built into transcript B (Fig. 1a). Next, the amiR-mP53/1- and $hRas^{G12V}$-expressing driver and the

(See figure on next page.)

Fig. 2 In vivo amiR-based gene silencing in the mouse liver. **a** Brightfield and fluorescence stereomicroscopic images of the liver of $Fah^{-/-}$ mice 5 months after the intrahepatic delivery of an amiR-free control and different amiR-expressing transposon vectors. **b** Monitoring the amount of the endogenous p53 mRNA and artificial transcripts A and B in the liver of $Fah^{-/-}$ mice 5 months after intrahepatic delivery of an amiR-free control and different amiR-expressing transposon vectors. Liver RNA samples were collected from $Fah^{-/-}$ mice at 5 months post-treatment. Samples were tested using Fah-, EGFP-, and p53 mRNA-specific RT-qPCR assays. Results were normalized to measurements of the ribosomal protein L27 (*Rpl27*) transcript as input control and data were presented as the mean ± SD ($n = 3$) (see Additional file 2 for individual data values and statistics). **c** Monitoring of endogenous p53 mRNA levels in NIH3T3 cells after stable transposon-based delivery of different amiR elements designed to silence *Tp53* expression. RNA samples were collected from cultured cells after G418 selection and tested using a p53 mRNA-specific RT-qPCR assay. Results were normalized to measurements of the ribosomal protein L27 (*Rpl27*) transcript as input control. Data were presented as the mean ± SD as relative values compared to the value generated using an amiR-free control vector ($n = 3$) (see Additional file 2 for individual data values and statistics). **d** Copy numbers of the transgenes in the liver of $Fah^{-/-}$ mice following intrahepatic delivery of different transposon vectors. Liver DNA samples were collected from $Fah^{-/-}$ mice at 5 months post-treatment. Samples were tested using a *Fah* transgene-specific qPCR assay. Results were normalized to measurements of the olfactory receptor 16 (*Olfr16*) gene as an input control, and values were presented relative to one diploid genome ($n = 3$) (see Additional file 2 for individual data values and statistics)

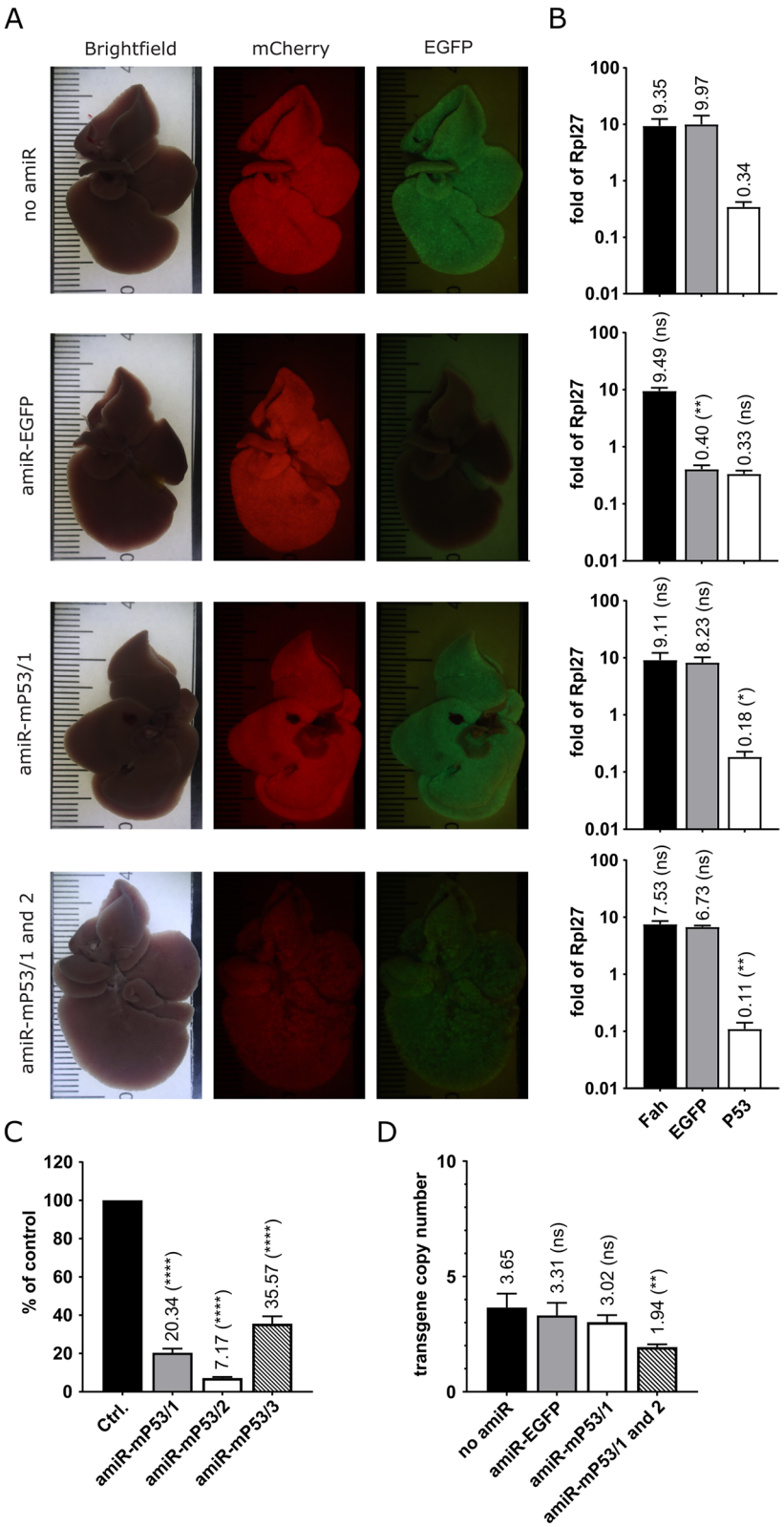

Fig. 2 (See legend on previous page.)

amiR-free EGFP-expressing control transposon constructs together with the SB100 transposase helper plasmid were co-delivered hydrodynamically into the liver of *Fah*$^{-/-}$ mice, then NTBC was withdrawn. After 5 weeks, during multi-nodular repopulation, the animals were sacrificed. Immunohistochemical investigations showed that the size of Fah-positive hepatocyte colonies repopulating the liver was larger in livers treated with the driver construct as compared to controls (Fig. 3a). This is consistent with the significantly higher proportion of Fah-positive hepatocytes (94.8% vs. 68.76%) in the driver construct-treated animals as determined by machine learning technology (Fig. 3b). This clearly shows that the division rate of the transformed hepatocytes is markedly higher than that of the regenerating cells in control mice during normal multi-nodular repopulation.

The examination time point at 5 weeks after hydrodynamic injection and NTBC withdrawal represents incomplete multi-nodular repopulation. At this time point, we again monitored the extent of *Tp53* silencing and the levels of transcripts A and B in both driver and control construct-treated animals. Our RT-qPCR measurements revealed 42% remaining *Tp53* expression in organs treated with the driver construct (Fig. 3c). The levels of transcripts A and B in the livers of animals treated with the control construct were very similar to each other (Fig. 3d), comparable to the findings in the later stages following complete multi-nodular repopulation (Figs. 1 c and 2b), whereas a significant decrease in the expression of transcript B (hRas^{G12V}) was observed in organs bearing the driver construct (Fig. 3d).

Experimental animals treated with the driver construct cannot be housed for substantially longer than 5 weeks due to the presence of a large number of transformed hepatocyte clones in their livers. Yet, to demonstrate the appearance of pathological signatures characteristic of tumors of malignant pathological grade using this high-penetrance driver combination, we mixed the transforming construct (1%) with a high amount of a transposon construct expressing the Fah selection marker protein alone (99%). Then, this construct mixture together with the SB100 transposase helper plasmid was co-delivered hydrodynamically into the liver of *Fah*$^{-/-}$ mice. Treated mice sacrificed at 5 months after NTBC withdrawal exhibited a high tumor burden (Additional file 1: Fig. S4). On histological images of their livers, tumor tissue showed high intrinsic fat accumulation typical for hepatocellular carcinoma (HCC) [24, 25] and frequent neoplastic tissue invasion and atypical mitoses were observed (Additional file 1: Fig. S4).

To better characterize the induced tumorigenic processes, immunohistochemical investigations were performed on liver samples from control and driver construct-treated animals. An early time point, 5 weeks post-treatment, and a late time point, 5 months post-treatment, were chosen (Fig. 3a). Samples from animals treated with the 1% transforming construct mixture were used to investigate the long-term effect of the driver construct at 5 months post-treatment. All liver samples from control and driver construct-treated animals were analyzed by immunohistochemical staining for alpha-fetoprotein (Afp), the most commonly used marker of HCC [26], and glypican-3 (Gpc3), one of its early markers [27, 28] (Fig. 3a). Liver samples from mice treated with the control construct did not show positivity for the Gpc3 marker at any of the stages tested (Fig. 3a). In contrast, weak positivity for the Afp marker was observed mainly in early-stage control samples at 5 weeks post-treatment (Fig. 3a). This weak Afp signal is presumably produced in remaining tyrosinemic cells [29, 30]. Interestingly, in the driver construct-treated samples, the majority of transformed hepatocyte colonies already showed Gpc3 positivity at the earlier time point of 5 weeks post-treatment, while Afp positivity was not observed at this stage. Mature tumors, 5 months after treatment, typically showed the presence of both markers tested (Fig. 3a).

(See figure on next page.)

Fig. 3 Induction of HCC using a predefined combination of drivers. **a** Immunohistochemical analysis of the Fah selection marker, Gpc3, and Afp HCC markers in liver sections from *Fah*$^{-/-}$ mice treated with either control (no amiR, EGFP) or driver (amiR-mP53/1, hRas^{G12V}) transposon constructs at 5 weeks and 5 months post-treatment. For the analysis of tumors emerging 5 months after treatment with the driver construct, a vector mixture containing 1% driver transposon vector and 99% transposon vector expressing only the Fah selection marker protein was used. Scale bars, 100 µm. **b** Determination of the percentage of Fah-positive hepatocytes 5 weeks after treatment by machine learning-based measurement. Data were presented as the mean ± SD ($n = 3$) (see Additional file 2 for individual data values and statistics). **c** Monitoring of endogenous p53 mRNA levels in the liver of *Fah*$^{-/-}$ mice treated with driver and control transposon constructs. Liver RNA samples were collected from *Fah*$^{-/-}$ mice at 5 weeks post-treatment and tested using a p53 mRNA-specific RT-qPCR assay. Results were normalized to measurements of the ribosomal protein L27 (*Rpl27*) transcript as input control and data were presented as the mean ± SD ($n = 3$) (see Additional file 2 for individual data values and statistics). **d** Monitoring the amount of transcripts A and B in the liver of *Fah*$^{-/-}$ mice treated with driver and control transposon constructs. Liver RNA samples were collected from *Fah*$^{-/-}$ mice at 5 weeks post-treatment and tested using Fah-, EGFP-, and hRas^{G12V}-specific RT-qPCR assays. Results were normalized to measurements of the ribosomal protein L27 (*Rpl27*) transcript as input control and data were presented as the mean ± SD ($n = 3$) (see Additional file 2 for individual data values and statistics)

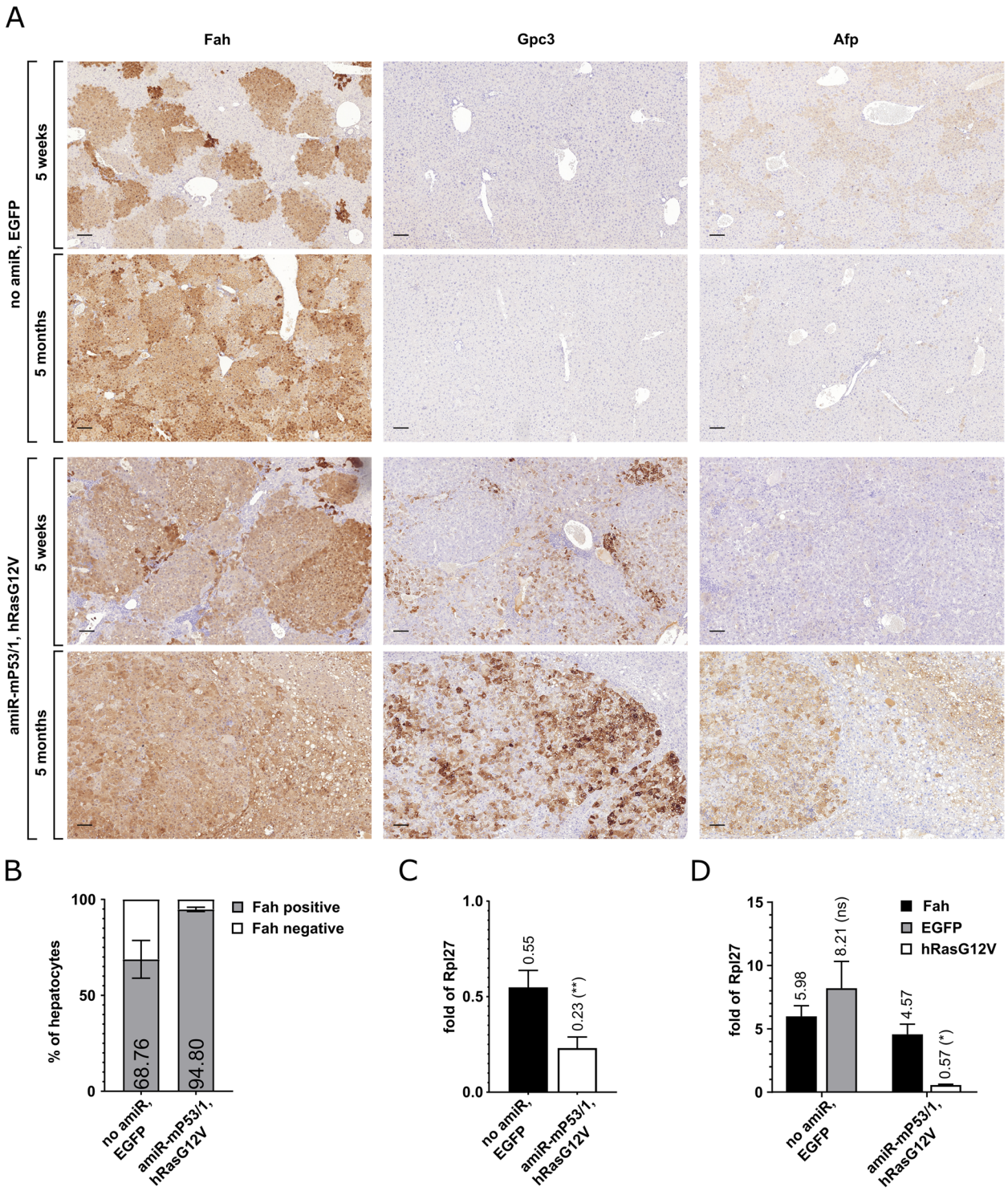

Fig. 3 (See legend on previous page.)

Discussion

The manipulation of the somatic genome in mice is hampered by a number of factors that are essentially the same as those that make it difficult to manipulate the human somatic genome for gene therapy. Transcriptional repression observed with plasmid DNA vectors in the liver is

caused by the formation of repressive heterochromatin on the plasmid DNA, a process initiated with the activation of cytoplasmic exogenous DNA sensors [8, 9]. Some studies suggest that heterochromatin formation occurs via CpG methylation [31, 32], whereas others propose a CpG-independent pathway [33]. As a consequence, transgene expression upon plasmid introduction, even if complemented by SB transposon-based chromosomal gene transfer rapidly declines in the liver [7, 10, 34]. Bell and co-workers demonstrated that even when using SB transposition, several weeks after injection, the transgene expression stabilizes at ~1% of the level at 24 h [10]. According to the results of our own comparative experiment presented here, Luc expression stabilizes at 0.26% of the level measured at day 3 in WT animals (Fig. 1e).

In comparison to the standard hydrodynamic injection method into WT animals, we can state that by exploiting the multi-nodular repopulation in *Fah*-deficient livers, a truly long-lasting and physiological level of gene expression can be achieved, virtually in the entire hepatocyte population of the experimental mice (Fig. 1b, d, e). In *Fah*-deficient animals, Luc expression showed increasing intensity over the 3-month monitoring period, up to 61-fold above the initial level (Fig. 1d, e). This net increase arises from several contributing factors, some acting in opposite directions. The initial activation of DNA sensors presumably cannot be avoided even with the technology described here. Therefore, the main negatively contributing factor is probably transgene heterochromatinization, which almost completely abolishes transgene expression in WT animals. We speculate that using our technology, the transposed transgenes are unlikely to undergo the degree of heterochromatinization required for silencing, probably due to positive selection pressure on Fah marker expression and to cell divisions during liver regeneration. The transient decrease in transgene expression in *Fah* KO animals at day 56 is presumably due to the fact that these downregulation processes are also at work in them, peaking at around day 56 but eventually ceasing (Fig. 1d, e). Another negative contributor is the disappearance of plasmid-derived transgene expression, which is due, among other things, to the loss of plasmids during cell divisions. Presumably, this phenomenon accounts for the lower initial (d3, d7) bioluminescence values measured in *Fah* KO animals (Fig. 1d, e). The main factor positively affecting cumulative transgene expression is the division of transgene-bearing cells. The 61-fold increase in transgene expression could be directly attributed to a similar increase in the number of hepatocytes, which corresponds to roughly 5 cell divisions, if negatively acting factors are excluded. However, as these are also present, more than 5 cell divisions are likely to occur during complete multi-nodular repopulation [12].

The SB transposon system has been used before for *Fah* correction [35]. Some laboratories have also taken significant steps forward to create a versatile gene expression modification system such as the one we present here [36–40]. The first study, and the one most similar to ours, was published by Wangensteen et al. [36]. The other similar studies typically use this vector system or its variants. The technological approach developed by Wangensteen et al. differs in some fundamental points from the one we describe here. The authors of the study used an early variant of amiR design [41] for endogenous gene silencing, which was later improved by others towards the mirE type structure [20]. Notably, this transcriptional unit expressing the early amiR variant was present on a separate transposon vector from the transposon vector expressing the Fah selection marker and the $NRas^{G12V}$ oncogene. Given that the toxicity of amiR elements may be significant [42, 43], target gene silencing was likely to affect only a subset of Fah-corrected $NRas^{G12V}$ expressing cells in regenerating liver tissue. Although this potentially low double feature positive cell count works with penetrant tumor models, it limits experimental applications and precludes settings in which all cells in a tissue need to be manipulated uniformly to draw valid conclusions. An artificial bidirectional promoter was also applied by Wangensteen et al. to connect the expression of the Fah selection marker and the $NRas^{G12V}$ oncogene, but the authors did not provide a detailed characterization of the amount of bidirectional transcripts relative to each other and to an endogenous control. A less effective liver repopulation as compared to our method can also be assumed potentially due to the application of less efficient SB transposase helper variants.

Here we used the hyperactive SB100 transposase helper for more effective chromosomal transgene delivery [11]. When designing our technology platform, we considered it important to allow tag-free expression of proteins of interest in such a way that their expression remains bound to the selection marker. The use of 2A-type peptides is not appropriate for this purpose, as they leave tags on both sides on the two transcriptionally linked proteins [17]. To overcome this challenge, we applied the HADHA/B endogenous bidirectional promoter capable of driving well-balanced bidirectional expression in the physiological range (Figs. 1 c and 2b). The use of HADHA/B gives the possibility for the marker-linked expression of an untagged native protein or a mutant protein isoform at the position of EGFP in transcript B (Fig. 1 a). Bidirectional promoters are more common than thought; a survey of the human genome indicated their widespread occurrence [44, 45]. The HADHA/B promoter is designed by nature to produce two subunits of a protein complex in stoichiometric proportions

[18]. According to our own measurements, the ratio of the two driven transcripts is indeed closer to 1/1 than for any other bidirectional promoter reported so far (Figs. 1c and 2b). In studies of either natural or synthetic bidirectional promoters, if such measurements have been performed at all, bidirectional transcript ratios worse than what we found for HADHA/B have been reported [46–50]. In a single case, we observed a significant shift from the typical 1/1 transcript A to B ratio to a ratio around 8/1 (Fig. 3d). Here the $hRas^{G12V}$ CDS was likely to exert negative selection pressure on the B side of the promoter, while positive pressure was still present on its A side (Fah), and in turn, the connection between the two sides of the promoter was weakened. The 1/1 transcript ratio produced by the HADHA/B bidirectional promoter can be particularly important for protein production when attempting to produce protein complexes.

In addition, we aimed to optimize amiR-based gene silencing in vivo in mouse liver. Importantly, we wanted this feature to be included in the same construct from which the other gene of interest (GOI) is expressed and on which the Fah selection marker is present. Thus, all genetic features are jointly represented in all Fah-corrected liver cells. We have also aimed to position the amiR element in such a way that its maturation does not interfere with the expression of either the selection marker or GOI. In light of this, we decided to place our amiR elements in an intron modified for this purpose. It is well known that in vivo stable gene silencing is not a straightforward technology. The in vivo toxicity of shRNA expression was long ago reported [42]. Similarly, amiR-expressing germline transgenic mice frequently show toxicity and viability issues. Miura and co-workers reported that it was not possible to generate either transgenic mice with higher amiR-expressing constructs or to generate homozygous mice with lower amiR-expressing constructs [43]. The potential reason was the saturation of endogenous miR-processing pathways or other non-specific toxic side effects of the applied amiRs. All these toxic side effects were enhanced with expression level. Here we did not observe severe, non-specific toxicity of gene silencing. In livers expressing the amiR-EGFP element, transcript A levels were virtually identical to those measured in controls not carrying amiR elements. While RT-qPCR measurements revealed only a slight non-significant decrease in levels of transcripts A and B in groups of animals carrying one or two amiR-mP53 elements as compared to controls (Fig. 2b), this slight decrease could be explained by negative selection against strong *Tp53* silencing, which is supported by the decrease in transgene copy number detected in these organs (Fig. 2d). Consequently, the physiological level of transgene expression provided by the human *HADHA/B* promoter in vivo in the mouse liver is compatible with the application of amiR elements.

No sign of tumorigenesis was detected in mice treated with EGFP-expressing transposon constructs either with or without intronic amiR structures. Even in strongly *Tp53* silenced mice carrying double amiR-mP53 elements, no tumor induction was observed. It is worth comparing these data with the published phenotype of *Tp53* KO mice [51]. Heterozygous *Tp53* KO animals rarely developed only lymphoid and testis tumors by 9 months of age, whereas approximately 75% of homozygous KO mice developed various tumors by 6 months of age. The most frequently observed tumors were malignant lymphomas, but no liver tumors were observed at all. Our results are in line with this, as we do not expect the development of hepatocellular carcinoma (HCC) in our experimental system within the given monitoring period even after aggressive *Tp53* silencing. This also implies that SB transposon-mediated random transgene delivery has no or very low oncogenic side effects, meaning that it is not sufficiently tumorigenic by itself to ruin our test system even in the presence of *Tp53* silencing.

Previously, it was shown that upon NTBC withdrawal a small subpopulation of Fah-deficient hepatocytes may be able to avoid elimination during the selection process by activating the survival Akt pathway [52]. Presumably, this phenomenon is responsible for the emergence of tumors from non-corrected Fah-deficient cells reported during retroviral gene therapy treatment of *Fah* KO mice [53]. In our system, such tumors of tyrosinemic cell origin are well distinguishable from marker-positive tumors induced by our transgenes, as they are negative for the mCherry marker. Their development requires the long-term presence of a large number of residual, uncorrected Fah-deficient cells in the liver. The emergence of this can be attributed to two potential causes: the poor efficiency of gene delivery and the use of contraselective transgenes. With the use of highly contraselective transgenes, during multi-nodular repopulation, the growth of nodules expressing these transgenes more strongly is retarded, slowing down the overall repopulation of the liver. Using the hyperactive SB100 transposase helper and non-contraselective transgenes, we do not see these tumors. However, the use of contraselective transgenes may also induce their appearance in our system in the long term. Such mCherry marker-negative tumors emerging from non-corrected Fah-deficient cells were only detected in double amiR-mP53 carrying mice at late time points, around 7–10 months post-injection. Consistent with the fact that tyrosinemic cells are prone to cancerous transformation, our measurements revealed that they might exhibit increased levels of *Tp53* expression as well. *Tp53* expression was measured in control (no amiR,

EGFP) animals at two time points, during multi-nodular repopulation, at 5 weeks after injection (Fig. 3c), and well after the completion of multi-nodular repopulation, at 5 months after injection (Fig. 2b). These RT-qPCR measurements revealed significantly higher *Tp53* expression levels in the first case (0.55-fold of Rpl27) at a considerable tyrosinemic (*Fah* KO) cell content than in the second case (0.34-fold of Rpl27), where tyrosinemic cells are already virtually absent.

Hepatocellular carcinoma (HCC) is one of the most lethal cancers worldwide; however, the genetic mechanisms underlying its pathogenesis are incompletely understood. We aimed to demonstrate the potential of our technology for cancer research by transforming mouse hepatocytes with the well-known *Tp53* tumor suppressor and *hRas* oncogene driver combination. Oncogenic hRas expression is able to trigger senescence in primary cells [54]. Therefore, their transformation by h*Ras* requires either a cooperating oncogene or the inactivation of a tumor suppressor [54, 55]. In accordance, ectopic expression of $hRas^{G12V}$ alone is insufficient to induce tumorigenesis. Thus, tumor development in our system alone demonstrates the combined manifestation of both cancer-driving genetic manipulations.

HCC characteristic of the induced tumors is well demonstrated by their Gpc3 and Afp immunoreactivity (Fig. 3a). It should also be noted that the presence of the Gpc3 marker, which can be used to distinguish early-stage HCC from dysplastic nodules [28], can already be detected in the majority of transformed hepatocyte colonies at the earlier examination time point, 5 weeks post-treatment (Fig. 3a).

The experimental potential created by our technology platform is not fully demonstrated by the *Tp53*-silenced and $hRas^{G12V}$-overexpressing tumor model. For the application of this highly penetrant driver combination, wild-type mice are also suitable, as it is sufficient to produce a small number of cells here that survive p53 silencing and overexpress $hRas^{G12V}$ to induce tumors. The high penetrance of oncogenic *Ras* mutations is a phenomenon parallel to the high incidence of *RAS* mutations in human cancer. Approximately 19% of patients with cancer harbor *RAS* mutations, equivalent to approximately 3.4 million new cases per year worldwide [56]. However, the current priority in cancer research is the functional validation of candidate driver genes with lower mutation frequencies in cancer genome databases. Such driver genes, when mutated, are likely to induce tumors at lower penetrance and may appear as germline or somatically mutated driver genes in hereditary as well as sporadic cancers [57]. The nearly 100 million testable hepatocytes available in our platform in a single experimental animal allow to functionally test even such driver genes with low-penetrance driver mutations.

Conclusions

In this study, we developed a somatic gene delivery technology enabling long-lasting and high-level transgene expression in the entire hepatocyte population of mice. We also presented comparative studies to demonstrate that our approach is superior to conventional methods. Our technology allows the tag-free expression of proteins of interest and silencing of any arbitrary gene in the mouse genome using amiR elements. Achieving these has been aided by the use of the endogenous HADHA/B promoter capable of driving well-balanced bidirectional expression and by optimizing in vivo intronic amiR-based gene silencing. The HADHA/B promoter has not been part of the biotechnology toolkit until now. Here we provide a detailed characterization of its functionality in an in vivo setting. Eventually, we developed a versatile technology platform for in vivo somatic genome editing in the mouse liver that simultaneously meets multiple requirements. We expect it will contribute to gene function analysis in mice by generating new, more accurate genetic models.

Methods

Animal care and maintain

Mice were bred and maintained in the Central Animal House at the Biological Research Centre (Szeged, Hungary). The specific pathogen-free status was confirmed quarterly according to FELASA (Federation for Laboratory Animal Science Associations) recommendations [58]. Mice were housed under 12-h light-dark cycle at 22 °C with free access to water and regular rodent chow. All animal experiments were conducted according to the protocols approved by the Institutional Animal Care and Use Committee at the Biological Research Centre. The used *Fah* mutant line, C57BL/6N-$Fah^{tm1(NCOM)Mfgc/Biat}$, is archived in the European Mouse Mutant Archive (EMMA) under EM:10787. $Fah^{-/-}$ mice were treated with 8 mg/l Orfadin®(Nitisinone, NTBC) (Swedish Orphan Biovitrum) in drinking water. After hydrodynamic injection, NTBC was withdrawn. C57BL/6NTac wild-type mice were obtained from Taconic Biosciences.

Plasmid construction

Empty pbiLiv-miR vector was synthesized and cloned in a pUC57 plasmid backbone by GeneScript. This encompasses the bidirectional promoter of the human hydroxyacyl-CoA dehydrogenase trifunctional multienzyme complex alpha (HADHA) and beta (HADHB) subunits. The HADHA side of the bidirectional promoter drives expression of the mCherry fluorescent marker gene,

which is disrupted by the first intron of the human eukaryotic translation elongation factor 1 alpha 1 (*EEF1A1*) to ensure intronic expression of the designed amiR structures. Restriction endonuclease recognition sites were introduced into the EEF1A1 intron to clone amiR elements as follows: AgeI, XbaI, SacI, and SalI. The mCherry coding sequence (CDS) is linked to the mouse fumaryl-aceto-acetate dehydrogenase (*Fah*) CDS by a T2A peptide to provide bicistronic expression. The transcription unit ends with a bGH polyadenylation signal. The HADHB side of the bidirectional promoter is flanked by an MCS followed by a bGH polyadenylation signal. The whole arrangement is flanked by the T2 type SB transposon inverted terminal repeats [14].

To generate pbiLiv-miR-EGFP, the PCR amplified EGFP coding sequence was inserted into the BamHI/PacI sites of the MCS in pbiLiv-miR. pbiLiv-miR-EGFP-EGFP and pbiLiv-miR-mP53/1-EGFP were constructed by inserting the complete amiR-EGFP or amiR-mp53/1 element into the AgeI/XbaI site of pbiLiv-miR-EGFP. pbiLiv-miR-mP53/1,2-EGFP was constructed by inserting the complete amiR-mP53/2 element into the SacI/SalI sites of pbiLiv-miR-mP53/1-EGFP. Complete amiR-EGFP, amiR-mp53/1, amiR-mp53/2, and amiR-mp53/3 elements (Additional file 1: Fig. S1) flanked by AgeI, SacI, XbaI, and SalI sites were synthesized and cloned in a pUC57 plasmid backbone by GeneScript.

To generate pbiLiv-miR-Luc, the Luciferase encoding gene was amplified by PCR from plasmid pGL3-Basic (Promega) and inserted into the NheI/PacI sites of the MCS in pbiLiv-miR.

For constructing pbiLiv-miR-mP53/1-hRas^{G12V}, hRas coding sequence was PCR amplified from mouse total liver RNA and cloned into the pBluescript SK plasmid. G12V mutation was introduced by the QuickChange Site Directed Mutagenesis Kit (Agilent Technologies). The mutated hRas coding sequence was then inserted into the BamHI/PacI sites of the MCS in pbiLiv-miR. Next, the amiR-mP53/1 element was inserted into the AgeI/XbaI sites of the hRas^{G12V}-containing pbiLiv-miR.

pNeo-miR was constructed by inserting the first intron of the *EEF1A1* with the AgeI, XbaI, SacI, and SalI restriction endonuclease recognition sites into pT2-SVNeo (Addgene #26553). The pNeo-miR-mP53 plasmid series was constructed by inserting the amiR-mP53/1, amiR-mP53/2, and amiR-mP53/3 amiR structures into the AgeI/XbaI sites of pNeo-miR. pcGlobin2-SB100 was constructed as described [11].

Hydrodynamic tail vein injection

Plasmids for hydrodynamic tail vein injection were prepared using the NucleoBond Xtra Maxi Plus EF Kit (Macherey-Nagel) according to the manufacturer's instructions. Before injection, we diluted plasmid DNA in Ringer's solution (0.9% NaCl, 0.03% KCl, 0.016% $CaCl_2$) and a volume equivalent to 10% of mouse body weight was administered via the lateral tail vein in 5–8 s into 6–8-week-old mice [10, 59]. The amount of plasmid DNA was 50 μg for each of the constructs mixed with 4 μg of the transposase helper plasmid.

RNA extraction and gene expression analysis

Total RNA from 50 mg liver tissue was isolated using TRI Reagent (MRC) following the manufacturer's protocol. RNA was Dnase I treated with PerfeCTa DNase I (Quantabio) and reverse transcribed into cDNA using RevertAid First Strand cDNA Synthesis Kit (ThermoFisher Scientific). RT-qPCR was performed on a Rotor-Gene Q instrument (Qiagen) with PerfeCTa SYBR Green SuperMix (Quantabio) as follows: 95 °C for 7 min followed by 35 cycles of 20 s at 95 °C, 20 s at 60 °C, and 20 s at 72 °C. All reactions were carried out in triplicates in a final volume of 20 μl. The following primers were used:

mP53-F: ACTTACCAGGGCAACTATGGCT; mP53-R: GCTGGCAGAATAGCTTATTGAGG; EGFP-F: GAGCAAAGACCCCAACGAGA; EGFP-R: CACCTCGAGCTACAGCTTCT; mFah-F: CTGGGTCAAGCTGCATGGAA; mFah-R: AGGAAGGTGCATTGTCGCAG; Rpl27-F: AAGCCGTCATCGTGAAGAACA; Rpl27-R: CTTGATCTTGGATCGCTTGGC [60].

PCR efficiencies were analyzed with Rotor-Gene Q software (Qiagen). Gene expression was analyzed by the normalization of expression to that of ribosomal protein L27 (*Rpl27*) using the ΔCT method [61].

Genomic DNA isolation and transgene copy number assessment

Whole livers of treated animals were lysed in 150 ml lysis buffer (100 mM TRIS-HCl pH8, 5 mM EDTA pH8, 200 mM NaCl, 0.2% SDS) and incubated overnight at 50 °C in the presence of 300 μg/ml ProteinaseK (VWR Chemicals). DNA from 1 ml lysate was isolated by conventional phenol/chloroform extraction and ethanol precipitation.

The assessment of transgene copy number was done by qPCR. Primers for Fah were the same as for the analysis of mRNA amount. PerfeCTa SYBR Green SuperMix (Quantabio) was used to carry out qPCR reactions on a Rotor-Gene Q instrument (Qiagen). All reactions were carried out in triplicates using 30 ng gDNA. Cycling conditions were as follows: 95 °C for 7 min followed by 35 cycles of 20 s at 95 °C, 20 s at 64 °C, and 20 s at 72 °C. PCR efficiencies were analyzed with Rotor-Gene Q software

(Qiagen). Relative changes in DNA levels were calculated using the ΔCT method. Results were normalized to measurements of the *Olfr16* gene. The following primers were used for this: Olfr16-F: GAGTTCGTCTTCCTGGGATTC; Olfr16-R: TAATGATGTTGCCAGCCAGA [62].

Stereomicroscope imaging

Pictures of whole mouse livers were taken with an Olympus SZX12 fluorescence stereozoom microscope equipped with a 100-W mercury lamp and filter sets for selective excitation and emission of GFP and mCherry.

In vivo bioluminescence imaging

In vivo bioluminescence imaging was performed using an IVIS Lumina III instrument (PerkinElmer). Following the intraperitoneal administration of 150 mg/kg luciferase substrate (XenoLight D-Luciferin-K+ salt, PerkinElmer), mice were anesthetized using isoflurane (Isoflurin, Vetpharma) and imaged at 10 min post-injection. Emitted photons were quantified with an exposure time of 1 to 10 s. Quantification of average radiance (photons per second per cm^2 per steradian (sr) [$p/s/cm^2/sr$]) within a circular region of interest was performed using the Living Image software (PerkinElmer).

Immunohistochemistry

Mice were sacrificed at 5 weeks, 3 months, or 5 months post-injection. Livers were removed and fixed overnight in 4% formalin, then embedded in paraffin, and cut into 5-µm sections. Immunohistochemistry was performed using the EnVision FLEX Mini Kit (DAKO). Antigen retrieval was done in a PT Link machine (DAKO). The primary antibodies used for immunohistochemistry are rabbit polyclonal anti-FAH (ThermoFisher Scientific, PA5-42049, 1:100), incubated for 120 min, and rabbit polyclonal anti-alpha 1 fetoprotein (Abcam, ab46799, 1:500) and rabbit polyclonal anti-glypican 3 (Abcam, ab186872, 1:800), incubated overnight. Secondary antibody polyclonal goat anti-rabbit-HRP (DAKO, P0448) was incubated for 30 min. Visualization was done with EnVision FLEX DAB+ Chromogen System (DAKO, GV825). After hematoxylin counterstaining for 5 min, slides were mounted and scanned with a Pannoramic Digital Slide Scanner (3D Histech).

Image analysis pipeline

3D Histech generated images were processed using BIAS software. A pipeline was created for the analysis consisting of four major steps: (1) pre-processing of the images, (2) segmentation and (3) feature extraction, and (4) cell classification using machine learning. In the pre-processing, non-uniform illumination was corrected using the CIDRE method [63]. The deep learning segmentation method was applied to detect and segment individual nuclei in images. With segmentation post-processing, two additional regions were defined for each nucleus: (1) a region representing the entire cell was defined by extending nuclei regions with a maximum 5 µm radius so that adjacent cells did not overlap and (2) cytoplasmic regions were defined by subtracting nuclei segmentation from the cell segmentation. Finally, morphological properties of these three different regions as well as intensity and texture features from all channels were extracted (in total 228 features) for cell classification. We employed supervised machine learning to predict four different cell types: FAH-positive cells, FAH-negative cells, immune cells, and other cells or segmentation artifacts that can be considered trash. These classes were manually selected based on their morphological characteristics. Cells with evenly distributed brown chromogen signal (anti-FAH staining) across the whole cells were labeled as FAH positive, while cells without chromogen staining were labeled as FAH negative. Cells with small and dark blue nuclei were considered as lymphocyte-like immune cells. Small segmented regions outside the tissue section were also classified as trash. For the training set, we annotated around 200 cells for each class from different tissue sections. Support Vector Machine (SVM) was trained with a radial basis function kernel commonly used for the multi-class cell phenotype classification. After training the SVM model, a 10-fold cross-validation was used to determine the expected accuracy of the model. We used this trained model to predict a class for all other cells in each liver section.

Protein extraction and Western blotting

Fifty milligrams of liver tissue was dounce homogenized in 2 ml radioimmunoprecipitation assay (RIPA) buffer (10 mmol/l Tris-HCl, pH 8.0, 1 mmol/l EDTA, 0.5 mmol/l EGTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 140 mmol/l NaCl), supplemented with PMSF (Merck). Cleared samples were sonicated for 3×10 s. Protein concentrations were calculated using the Pierce™ BCA Protein Assay kit (ThermoFisher Scientific). A total of 80 µg of protein was separated on 10% SDS-PAGE gel, transferred to a 0.2-µm nitrocellulose membrane (Amersham), and blocked with 5% non-fat dry milk in Tris-buffered saline-Tween 20 (TBS-T) for 1 h at room temperature. Blocked membranes were incubated with anti-GFP (Abcam, ab6556, 1:4000) and peroxidase-conjugated anti-GAPDH (ThermoFisher Scientific, MA5-15738-HRP,1:10000) antibodies. Anti-rabbit IgG conjugated to horseradish peroxidase (HRP) (Sigma Aldrich, A0545, 1:20000) was used as the secondary antibody where necessary. Immune complexes detected with

enhanced chemiluminescence (ECL) Prime Western blotting Detection Reagent (Amersham).

Cell culture and transfection

NIH/3T3 cells were purchased from ATCC and cultured in Dulbecco's modified Eagle's medium (DMEM) (Biosera) supplemented with 10% fetal bovine serum (FBS; Gibco) and 1% penicillin-streptomycin (P-S; HyClone) in the presence of 5% CO_2 at 37 °C.

Cells were transfected with 500 ng of either pNeo-miR or pNeo-miR-mP53/1 or pNeo-miR-mP53/2 or pNeo-miR-mP53 in combination with 50 ng of the transposase helper plasmid, using FuGENE® HD transfection reagent (Promega) according to manufacturer's instructions. Selection of stable transfected cells was performed using neomycin (G418; Biosera).

Data visualization and statistics

GraphPad Prism software (version 8.4.3 for Windows, GraphPad Software) was used for data visualization. Statistics were calculated using Fisher's exact and Pearson's chi-squared test.

To identify levels of statistical significance (*P* value), one-way ANOVA tests were performed for the comparison of RT-qPCR measurements of different sample groups and for the comparison of two sample groups the Welch's *t*-test was applied. The threshold for significance was $P<0.05$.

Abbreviations

Afp	Alpha-fetoprotein
AIM2	Absent in melanoma 2
amiR	Artificial microRNA
CDS	Coding sequence
cGAS	Cyclic GMP-AMP synthase
EEF1A1	Eukaryotic translation elongation factor 1 alpha 1
EGFP	Enhanced green fluorescent protein
Fah	Fumarylacetoacetate hydrolase
GOI	Gene of interest
Gpc3	Glypican-3
HADHA/B	Hydroxyacyl-CoA dehydrogenase trifunctional multienzyme complex alpha and beta subunits
HCC	Hepatocellular carcinoma
IHC	Immunohistochemistry
ITR	Inverted terminal repeat
KO	Knock out
Luc	Firefly luciferase
miR	MicroRNA
NTBC	2-(2-Nitro-4-trifluoromethylbenzoyl)-1,3-cyclohexanedione (Orfadin®, Nitisinon)
Rpl27	Ribosomal protein 27
RT-qPCR	Quantitative reverse transcription PCR
SB	Sleeping Beauty transposon
shRNA	Short hairpin RNA
TLR9	Toll-like receptor 9
WT	Wild type

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12915-022-01262-x.

Additional file 1: Fig. S1. The structure of the applied amiR elements. **Fig. S2.** Detection of the EGFP protein in mouse liver. **Fig. S3.** SB transposon-based cloning platform for the expression of amiR elements in cultured cells. **Fig. S4.** Stereomicroscopic and histological examination of liver samples 5 months after treatment with a construct mixture containing 1% transforming construct.

Additional file 2. Individual data values and statistics for Figures 1-3.

Acknowledgements

The authors thank the staff of the Cellular Imaging Laboratory of the Biological Research Centre for the provided microscopy support.

Consent to participate

Not applicable.

Authors' contributions

LM contributed to the conception and design. LM and AN contributed to the writing, review, and revision of the manuscript with input from all authors. LM contributed to the analysis and interpretation of data with help from IO, LGP, PH, and FS. RK, LH, and GI performed the animal experiments. GI performed the IVIS Lumina imaging. AGK, DZP, IF, and VED participated in the construction and purification of plasmids. AGK, DZP, PG, ABD, and AM carried out qPCR and RT-qPCR experiments. GPB and KSAA performed the immunohistochemistry. EM and AK performed the image analysis. DZP, VED, and AGK carried out the cell culture experiments. AGK and RK performed the Western blot experiment. LM contributed to the study supervision. All authors read and approved the final manuscript.

Funding

This work was supported by the Momentum Programme of the Hungarian Academy of Sciences [LP2015-5/2015] and by the National Research, Development and Innovation Office (Hungary) grant [GINOP-2.3.2-15-2016-00024]. PH, EM, and AK acknowledge support from the LENDULET-BIOMAG Grant [2018-342], from the European Regional Development Fund (GINOP-2.3.2-15-2016-00026), from COMPASS-ERA PerMed H2020, from CZI Deep Visual Proteomics, from H2020-DiscovAir, and from ELKH-Excellence grant. RK was supported by the ÚNKP-20-3-SZTE-84 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund. AGK was supported by the Szeged Scientists Academy under the sponsorship of the Hungarian Ministry of Innovation and Technology (*FEIF/646-4/2021-ITM_SZERZ*) and the New National Excellence Program of the Ministry for Innovation and Technology of Hungary (*ÚNKP-20-2 –SZTE-438*). Open access funding provided by ELKH Biological Research Center.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and supplementary information files. Individual data values and statistics are provided in Additional file 2. Presented qPCR experiments comply with the MIQE Guidelines. Necessary details for evaluation are supplied in the "Methods" section of the manuscript. The exact sequences of artificial microRNA structures used are provided in Additional file 1: Fig. S1.

Declarations

Ethics approval

Animal experimental procedures were approved and performed in accordance with the guidelines of the Institutional Animal Care and Use Committee at the Biological Research Centre (Szeged, Hungary) under the supervision of the Governmental Office for Csongrád County, Directorate of Food Chain Safety and Animal Health. The approval number is XVI./801/2018.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Institute of Genetics, Biological Research Centre, Szeged, Hungary. [2]Doctoral School of Biology, University of Szeged, Szeged, Hungary. [3]Doctoral School of Multidisciplinary Medical Sciences, University of Szeged, Szeged, Hungary. [4]Institute of Pathology, University of Szeged, Szeged, Hungary. [5]Synthetic and Systems Biology Unit, Institute of Biochemistry, Biological Research Centre, Szeged, Hungary. [6]Department of Biochemistry, University of Szeged, Szeged, Hungary. [7]Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland.

Received: 25 August 2021 Accepted: 22 February 2022
Published online: 01 April 2022

References

1. Whitelaw CB, Springbett AJ, Webster J, Clark J. The majority of G0 transgenic mice are derived from mosaic embryos. Transgenic Res. 1993;2(1):29–32. https://doi.org/10.1007/BF01977678 PubMed PMID: 8513336.
2. Garrick D, Fiering S, Martin DI, Whitelaw E. Repeat-induced gene silencing in mammals. Nat Genet. 1998;18(1):56–9. https://doi.org/10.1038/ng0198-56 PubMed PMID: 9425901.
3. Limberis MP, Bell CL, Heath J, Wilson JM. Activation of transgene-specific T cells following lentivirus-mediated gene delivery to mouse lung. Mol Ther. 2010;18(1):143–50. https://doi.org/10.1038/mt.2009.190 PubMed PMID: 19724265; PubMed Central PMCID: PMCPMC2839217.
4. Zhu J, Huang X, Yang Y. Innate immune response to adenoviral vectors is mediated by both Toll-like receptor-dependent and -independent pathways. J Virol. 2007;81(7):3170–80. https://doi.org/10.1128/JVI.02192-06 PubMed PMID: 17229689; PubMed Central PMCID: PMCPMC1866082.
5. Liu L, Zern MA, Lizarzaburu ME, Nantz MH, Wu J. Poly (cationic lipid)-mediated in vivo gene delivery to mouse liver. Gene Ther. 2003;10(2):180–7. https://doi.org/10.1038/sj.gt.3301861 PubMed PMID: 12571647.
6. Lecocq M, Andrianaivo F, Warnier MT, Wattiaux-De Coninck S, Wattiaux R, Jadot M. Uptake by mouse liver and intracellular fate of plasmid DNA after a rapid tail vein injection of a small or a large volume. J Gene Med. 2003;5(2):142–56. https://doi.org/10.1002/jgm.328 PubMed PMID: 12539152.
7. Herweijer H, Zhang G, Subbotin VM, Budker V, Williams P, Wolff JA. Time course of gene expression after plasmid DNA gene transfer to the liver. J Gene Med. 2001;3(3):280–91. https://doi.org/10.1002/jgm.178 PubMed PMID: 11437333.
8. Broz P, Dixit VM. Inflammasomes: mechanism of assembly, regulation and signalling. Nat Rev Immunol. 2016;16(7):407–20. https://doi.org/10.1038/nri.2016.58 PubMed PMID: 27291964.
9. Pandey S, Kawai T, Akira S. Microbial sensing by Toll-like receptors and intracellular nucleic acid sensors. Cold Spring Harb Perspect Biol. 2014;7(1):a016246. https://doi.org/10.1101/cshperspect.a016246 PubMed PMID: 25301932; PubMed Central PMCID: PMCPMC4292165.
10. Bell JB, Podetz-Pedersen KM, Aronovich EL, Belur LR, McIvor RS, Hackett PB. Preferential delivery of the Sleeping Beauty transposon system to livers of mice by hydrodynamic injection. Nat Protoc. 2007;2(12):3153–65. https://doi.org/10.1038/nprot.2007.471 PubMed PMID: 18079715; PubMed Central PMCID: PMCPMC2548418.
11. Mates L, Chuah MK, Belay E, Jerchow B, Manoj N, Acosta-Sanchez A, et al. Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. Nat Genet. 2009;41(6):753–61. https://doi.org/10.1038/ng.343ng.343 Epub 2009/05/05. [pii]. PubMed PMID: 19412179.
12. Overturf K, Al-Dhalimy M, Tanguay R, Brantly M, Ou CN, Finegold M, et al. Hepatocytes corrected by gene therapy are selected in vivo in a murine model of hereditary tyrosinaemia type I. Nat Genet. 1996;12(3):266–73. https://doi.org/10.1038/ng0396-266 PubMed PMID: 8589717.
13. Zhang G, Gao X, Song YK, Vollmer R, Stolz DB, Gasiorowski JZ, et al. Hydroporation as the mechanism of hydrodynamic delivery. Gene Ther. 2004;11(8):675–82. https://doi.org/10.1038/sj.gt.3302210 PubMed PMID: 14724673; PubMed Central PMCID: PMCPMC4412368.
14. Cui Z, Geurts AM, Liu G, Kaufman CD, Hackett PB. Structure-function analysis of the inverted terminal repeats of the sleeping beauty transposon. J Mol Biol. 2002;318(5):1221–35. https://doi.org/10.1016/s0022-2836(02)00237-1 PubMed PMID: 12083513.
15. Skipper KA, Andersen PR, Sharma N, Mikkelsen JG. DNA transposon-based gene vehicles - scenes from an evolutionary drive. J Biomed Sci. 2013;20:92. https://doi.org/10.1186/1423-0127-20-92 PubMed PMID: 24320156; PubMed Central PMCID: PMCPMC3878927.
16. Lehmann K, Tschuor C, Rickenbacher A, Jang JH, Oberkofler CE, Tschopp O, et al. Liver failure after extended hepatectomy in mice is mediated by a p21-dependent barrier to liver regeneration. Gastroenterology. 2012;143(6):1609–19 e4. https://doi.org/10.1053/j.gastro.2012.08.043 PubMed PMID: 22960658.
17. Szymczak AL, Vignali DA. Development of 2A peptide-based strategies in the design of multicistronic vectors. Expert Opin Biol Ther. 2005;5(5):627–38. https://doi.org/10.1517/14712598.5.5.627 PubMed PMID: 15934839.
18. Orii KE, Orii KO, Souri M, Orii T, Kondo N, Hashimoto T, et al. Genes for the human mitochondrial trifunctional protein alpha- and beta-subunits are divergently transcribed from a common promoter region. J Biol Chem. 1999;274(12):8077–84. https://doi.org/10.1074/jbc.274.12.8077 PubMed PMID: 10075708.
19. Beisel CL, Chen YY, Culler SJ, Hoff KG, Smolke CD. Design of small molecule-responsive microRNAs based on structural requirements for Drosha processing. Nucleic Acids Res. 2011;39(7):2981–94. https://doi.org/10.1093/nar/gkq954 PubMed PMID: 21149259; PubMed Central PMCID: PMCPMC3074164.
20. Fellmann C, Hoffmann T, Sridhar V, Hopfgartner B, Muhar M, Roth M, et al. An optimized microRNA backbone for effective single-copy RNAi. Cell Rep. 2013;5(6):1704–13. https://doi.org/10.1016/j.celrep.2013.11.020 PubMed PMID: 24332856.
21. Dow LE, Premsrirut PK, Zuber J, Fellmann C, McJunkin K, Miething C, et al. A pipeline for the generation of shRNA transgenic mice. Nat Protoc. 2012;7(2):374–93. https://doi.org/10.1038/nprot.2011.446 PubMed PMID: 22301776; PubMed Central PMCID: PMCPMC3724521.
22. Baratta JL, Ngo A, Lopez B, Kasabwalla N, Longmuir KJ, Robertson RT. Cellular organization of normal mouse liver: a histological, quantitative immunocytochemical, and fine structural analysis. Histochem Cell Biol. 2009;131(6):713–26. https://doi.org/10.1007/s00418-009-0577-1 PubMed PMID: 19255771; PubMed Central PMCID: PMCPMC2761764.
23. Bos JL. The ras gene family and human carcinogenesis. Mutat Res. 1988;195(3):255–71. https://doi.org/10.1016/0165-1110(88)90004-8 PubMed PMID: 3283542.
24. Berndt N, Eckstein J, Heucke N, Gajowski R, Stockmann M, Meierhofer D, et al. Characterization of lipid and lipid droplet metabolism in human HCC. Cells. 2019;8(5). https://doi.org/10.3390/cells8050512 PubMed PMID: 31137921; PubMed Central PMCID: PMCPMC6562484.
25. Currie E, Schulze A, Zechner R, Walther TC, Farese RV Jr. Cellular fatty acid metabolism and cancer. Cell Metab. 2013;18(2):153–61. https://doi.org/10.1016/j.cmet.2013.05.017 PubMed PMID: 23791484; PubMed Central PMCID: PMCPMC3742569.
26. Gorog D, Regoly-Merei J, Paku S, Kopper L, Nagy P. Alpha-fetoprotein expression is a potential prognostic marker in hepatocellular carcinoma. World J Gastroenterol. 2005;11(32):5015–8. https://doi.org/10.3748/wjg.v11.i32.5015 PdubMed PMID: 16124056; PubMed Central PMCID: PMCPMC4321920.
27. Anatelli F, Chuang ST, Yang XJ, Wang HL. Value of glypican 3 immunostaining in the diagnosis of hepatocellular carcinoma on needle biopsy. Am J Clin Pathol. 2008;130(2):219–23. https://doi.org/10.1309/WMB5PX57Y4P8QCTY PubMed PMID: 18628090.
28. Llovet JM, Chen Y, Wurmbach E, Roayaie S, Fiel MI, Schwartz M, et al. A molecular signature to discriminate dysplastic nodules from early hepatocellular carcinoma in HCV cirrhosis. Gastroenterology. 2006;131(6):1758–67. https://doi.org/10.1053/j.gastro.2006.09.014 PubMed PMID: 17087938.
29. Angileri F, Roy V, Morrow G, Scoazec JY, Gadot N, Orejuela D, et al. Molecular changes associated with chronic liver damage and neoplastic lesions in a murine model of hereditary tyrosinemia type 1. Biochim Biophys

Acta. 2015;1852(12):2603–17. https://doi.org/10.1016/j.bbadis.2015.09.002 PubMed PMID: 26360553.
30. Zerbini C, Weinberg DS, Hollister KA, Perez-Atayde AR. DNA ploidy abnormalities in the liver of children with hereditary tyrosinemia type I. Correlation with histopathologic features. Am J Pathol. 1992;140(5):1111–9 PubMed PMID: 1374592; PubMed Central PMCID: PMCPMC1886502.
31. Yew NS, Zhao H, Przybylska M, Wu IH, Tousignant JD, Scheule RK, et al. CpG-depleted plasmid DNA vectors with enhanced safety and long-term gene expression in vivo. Mol Ther. 2002;5(6):731–8. https://doi.org/10.1006/mthe.2002.0598 PubMed PMID: 12027557.
32. Hodges BL, Taylor KM, Joseph MF, Bourgeois SA, Scheule RK. Long-term transgene expression from plasmid DNA gene therapy vectors is negatively affected by CpG dinucleotides. Mol Ther. 2004;10(2):269–78. https://doi.org/10.1016/j.ymthe.2004.04.018 PubMed PMID: 15294174.
33. Chen ZY, Riu E, He CY, Xu H, Kay MA. Silencing of episomal transgene expression in liver by plasmid bacterial backbone DNA is independent of CpG methylation. Mol Ther. 2008;16(3):548–56. https://doi.org/10.1038/sj.mt.6300399 PubMed PMID: 18253155.
34. Bell JB, Aronovich EL, Schreifels JM, Beadnell TC, Hackett PB. Duration of expression and activity of Sleeping Beauty transposase in mouse liver following hydrodynamic DNA delivery. Mol Ther. 2010;18(10):1796–802. https://doi.org/10.1038/mt.2010.152 PubMed PMID: 20628359; PubMed Central PMCID: PMCPMC2951564.
35. Montini E, Held PK, Noll M, Morcinek N, Al-Dhalimy M, Finegold M, et al. In vivo correction of murine tyrosinemia type I by DNA-mediated transposition. Mol Ther. 2002;6(6):759–69. https://doi.org/10.1006/mthe.2002.0812 PubMed PMID: 12498772.
36. Wangensteen KJ, Wilber A, Keng VW, He Z, Matise I, Wangensteen L, et al. A facile method for somatic, lifelong manipulation of multiple genes in the mouse liver. Hepatology. 2008;47(5):1714–24. https://doi.org/10.1002/hep.22195 PubMed PMID: 18435462; PubMed Central PMCID: PMCPMC5808937.
37. Keng VW, Tschida BR, Bell JB, Largaespada DA. Modeling hepatitis B virus X-induced hepatocellular carcinoma in mice with the Sleeping Beauty transposon system. Hepatology. 2011;53(3):781–90. https://doi.org/10.1002/hep.24091 PubMed PMID: 21374658; PubMed Central PMCID: PMCPMC3079950.
38. Riordan JD, Keng VW, Tschida BR, Scheetz TE, Bell JB, Podetz-Pedersen KM, et al. Identification of rtl1, a retrotransposon-derived imprinted gene, as a novel driver of hepatocarcinogenesis. PLoS Genet. 2013;9(4):e1003441. https://doi.org/10.1371/journal.pgen.1003441 PubMed PMID: 23593033; PubMed Central PMCID: PMCPMC3616914.
39. Wuestefeld T, Pesic M, Rudalska R, Dauch D, Longerich T, Kang TW, et al. A Direct in vivo RNAi screen identifies MKK4 as a key regulator of liver regeneration. Cell. 2013;153(2):389–401. https://doi.org/10.1016/j.cell.2013.03.026 PubMed PMID: 23582328.
40. Tschida BR, Temiz NA, Kuka TP, Lee LA, Riordan JD, Tierrablanca CA, et al. Sleeping Beauty insertional mutagenesis in mice identifies drivers of steatosis-associated hepatic tumors. Cancer Res. 2017;77(23):6576–88. https://doi.org/10.1158/0008-5472.CAN-17-2281 PubMed PMID: 28993411; PubMed Central PMCID: PMCPMC5712258.
41. Dickins RA, Hemann MT, Zilfou JT, Simpson DR, Ibarra I, Hannon GJ, et al. Probing tumor phenotypes using stable and regulated synthetic microRNA precursors. Nat Genet. 2005;37(11):1289–95. https://doi.org/10.1038/ng1651 PubMed PMID: 16200064.
42. Grimm D, Streetz KL, Jopling CL, Storm TA, Pandey K, Davis CR, et al. Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. Nature. 2006;441(7092):537–41. https://doi.org/10.1038/nature04791 PubMed PMID: 16724069.
43. Miura H, Inoko H, Tanaka M, Nakaoka H, Kimura M, Gurumurthy CB, et al. Assessment of artificial MiRNA architectures for higher knockdown efficiencies without the undesired effects in mice. PLoS One. 2015;10(8):e0135919. https://doi.org/10.1371/journal.pone.0135919 PubMed PMID: 26285215; PubMed Central PMCID: PMCPMC4540464.
44. Adachi N, Lieber MR. Bidirectional gene organization: a common architectural feature of the human genome. Cell. 2002;109(7):807–9. https://doi.org/10.1016/s0092-8674(02)00758-4 PubMed PMID: 12110178.
45. Trinklein ND, Aldred SF, Hartman SJ, Schroeder DI, Otillar RP, Myers RM. An abundance of bidirectional promoters in the human genome. Genome Res. 2004;14(1):62–6. https://doi.org/10.1101/gr.1982804 PubMed PMID: 14707170; PubMed Central PMCID: PMCPMC314279.
46. Amendola M, Venneri MA, Biffi A, Vigna E, Naldini L. Coordinate dual-gene transgenesis by lentiviral vectors carrying synthetic bidirectional promoters. Nat Biotechnol. 2005;23(1):108–16. https://doi.org/10.1038/nbt1049 PubMed PMID: 15619618.
47. He K, Rad S, Poudel A, McLellan AD. Compact bidirectional promoters for dual-gene expression in a Sleeping Beauty transposon. Int J Mol Sci. 2020;21(23). https://doi.org/10.3390/ijms21239256 PubMed PMID: 33291599; PubMed Central PMCID: PMCPMC7731152.
48. Andrianaki A, Siapati EK, Hirata RK, Russell DW, Vassilopoulos G. Dual transgene expression by foamy virus vectors carrying an endogenous bidirectional promoter. Gene Ther. 2010;17(3):380–8. https://doi.org/10.1038/gt.2009.147 PubMed PMID: 19907502; PubMed Central PMCID: PMCPMC3739712.
49. Golding MC, Mann MR. A bidirectional promoter architecture enhances lentiviral transgenesis in embryonic and extraembryonic stem cells. Gene Ther. 2011;18(8):817–26. https://doi.org/10.1038/gt.2011.26 PubMed PMID: 21390068.
50. Na M, Fan X. Design of Ad5F35 vectors for coordinated dual gene expression in candidate human hematopoietic stem cells. Exp Hematol. 2010;38(6):446–52. https://doi.org/10.1016/j.exphem.2010.03.007 PubMed PMID: 20303383.
51. Donehower LA, Harvey M, Slagle BL, McArthur MJ, Montgomery CA Jr, Butel JS, et al. Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. Nature. 1992;356(6366):215–21. https://doi.org/10.1038/356215a0 PubMed PMID: 1552940.
52. Orejuela D, Jorquera R, Bergeron A, Finegold MJ, Tanguay RM. Hepatic stress in hereditary tyrosinemia type 1 (HT1) activates the AKT survival pathway in the fah-/- knockout mice model. J Hepatol. 2008;48(2):308–17. https://doi.org/10.1016/j.jhep.2007.09.014 PubMed PMID: 18093685.
53. Grompe M, Overturf K, Al-Dhalimy M, Finegold M. Therapeutic trials in the murine model of hereditary tyrosinaemia type I: a progress report. J Inherit Metab Dis. 1998;21(5):518–31. https://doi.org/10.1023/a:1005462804271 PubMed PMID: 9728332.
54. Serrano M, Lin AW, McCurrach ME, Beach D, Lowe SW. Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. Cell. 1997;88(5):593–602. https://doi.org/10.1016/s0092-8674(00)81902-9 PubMed PMID: 9054499.
55. Ju HL, Ahn SH, Kim DY, Baek S, Chung SI, Seong J, et al. Investigation of oncogenic cooperation in simple liver-specific transgenic mouse models using noninvasive in vivo imaging. PLoS One. 2013;8(3):e59869. https://doi.org/10.1371/journal.pone.0059869 PubMed PMID: 23555816; PubMed Central PMCID: PMCPMC3610734.
56. Prior IA, Hood FE, Hartley JL. The frequency of Ras mutations in cancer. Cancer Res. 2020;80(14):2969–74. https://doi.org/10.1158/0008-5472.CAN-19-3682 PubMed PMID: 32209560; PubMed Central PMCID: PMCPMC7367715.
57. Gohler S, Da Silva Filho MI, Johansson R, Enquist-Olsson K, Henriksson R, Hemminki K, et al. Functional germline variants in driver genes of breast cancer. Cancer Causes Control. 2017;28(4):259–71. https://doi.org/10.1007/s10552-017-0849-3 PubMed PMID: 28238063.
58. Mahler Convenor M, Berard M, Feinstein R, Gallagher A, Illgen-Wilcke B, Pritchett-Corning K, et al. FELASA recommendations for the health monitoring of mouse, rat, hamster, guinea pig and rabbit colonies in breeding and experimental units. Lab Anim. 2014;48(3):178–92 PubMed PMID: Medline:24496575.
59. Portier I, Vanhoorelbeke K, Verhenne S, Pareyn I, Vandeputte N, Deckmyn H, et al. High and long-term von Willebrand factor expression after Sleeping Beauty transposon-mediated gene therapy in a mouse model of severe von Willebrand disease. J Thromb Haemost. 2018;16(3):592–604. https://doi.org/10.1111/jth.13938 PubMed PMID: 29288565.
60. Thomas KC, Zheng XF, Garces Suarez F, Raftery JM, Quinlan KG, Yang N, et al. Evidence based selection of commonly used RT-qPCR reference genes for the analysis of mouse skeletal muscle. PLoS One. 2014;9(2):e88653. https://doi.org/10.1371/journal.pone.0088653 PubMed PMID: 24523926; PubMed Central PMCID: PMCPMC3921188.
61. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods.

2001;25(4):402–8. https://doi.org/10.1006/meth.2001.1262 PubMed PMID: 11846609.
62. D'Hulst C, Parvanova I, Tomoiaga D, Sapar ML, Feinstein P. Fast quantitative real-time PCR-based screening for common chromosomal aneuploidies in mouse embryonic stem cells. Stem Cell Reports. 2013;1(4):350–9. https://doi.org/10.1016/j.stemcr.2013.08.003 PubMed PMID: 24319669; PubMed Central PMCID: PMCPMC3849352.
63. Smith K, Li Y, Piccinini F, Csucs G, Balazs C, Bevilacqua A, et al. CIDRE: an illumination-correction method for optical microscopy. Nat Methods. 2015;12(5):404–6. https://doi.org/10.1038/nmeth.3323 PubMed PMID: 25775044.

Publisher's Note

Co-author certification

I, myself as a corresponding author of the following publication(s) declare that the authors have no conflict of interest, and **Réka Karkas**, Ph.D. candidate had significant contribution to the jointly published research(es). The results discussed in her thesis were not used and not intended to be used in any other qualification process for obtaining a PhD degree.

Szeged, 2025.03.07.

.....

corresponding author

The publication(s) relevant to the applicant's thesis:

I. **Réka Karkas**, Khaloon Sadiq Ahmed Abdullah*, Laszlo Kaizer, Ádám Ürmös, Lilla Tiszlavicz, Tibor Pankotai, Istvan Nagy, Lajos Mátés, Farkas Sükösd: *LINE-1 ORF1p is a Promising Biomarker in Cervical Intraepithelial Neoplasia Degree Assessment* International. Journal of Gynecological Pathology (2024) DOI: 10.1097/PGP.0000000000001035 (Q2)

Dr. Réka Karkas participated in patient data acquisition and FFPE sample collection and immunostaining. She participated in the development of the modified version of the immunoreactive scale (IRS) to evaluate ORF1p immunostaining results. She participated in the evaluation of the experimental data.

II. Anna Georgina Kopasz, Dávid Zsolt Pusztai, **Réka Karkas**, Liza Hudoba, Khaldoon Sadiq Ahmed Abdullah, Gergely Imre, Gabriella Pankotai-Bodó, Ede Migh, Andrea Nagy, András Kriston, Péter Germán, Andrea Bakné Drubi, Anna Molnár, Ildikó Fekete, Virág Éva Dani, Imre Ocsovszki, László Géza Puskás, Péter Horváth, Farkas Sükösd, Lajos Mátés: *A versatile transposon-based technology to generate loss- and gain-of-function phenotypes in the mouse liver* BMC BIOLOGY 20 : 1 Paper: 74 , 17 p. (2022) BMC Biology 20, 74 (2022). DOI: 10.1186/s12915-022-01262-x (D1)

Dr. Réka Karkas participated in the Hydrodynamic tail vein injections, RNA extraction, Genomic DNA isolation, protein isolation, Western blotting, in vivo bioluminescence imaging and Immunohistochemistry. She participated in the evaluation of the experimental data.