



The human EF1a promoter does not provide expression of the transgene in mice

Nariman Battulin[✉] · Alexey Korablev · Anastasia Ryzhkova · Alexander Smirnov · Evelyn Kabirova · Anna Khabarova · Timofey Lagunov · Irina Serova · Oleg Serov

Received: 16 May 2022 / Accepted: 3 August 2022 / Published online: 12 August 2022
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

Abstract In this work, we set out to create mice susceptible to the SARS-CoV-2 coronavirus. To ensure the ubiquitous expression of the human ACE2 gene we used the human EF1a promoter. Using pronuclear microinjection of the transgene construct, we obtained six founders with the insertion of the EF1a-hACE2 transgene, from which four independent mouse lines were established. Unfortunately, only one line had low levels of hACE2 expression in some organs. In addition, we did not detect the hACE2 protein in primary lung fibroblasts from any of the transgenic lines. Bisulfite sequencing analysis revealed that the EF1a promoter was hypermethylated in the genomes of transgenic animals. Extensive analysis of published works about transgenic animals indicated that EF1a transgenic constructs are frequently inactive. Thus, our case cautions against using the EF1a

promoter to generate transgenic animals, as it is prone to epigenetic silencing.

Keywords EF1a promoter · Transgenesis · Pronuclear microinjection · DNA methylation · Expression

Introduction

Transgenic animals are a powerful tool for the study of many biological processes. The exact activity of the transgene depends on the promoter chosen, so the choice of promoter is one of the critical decisions when designing a genetic construct. For many tasks, it is essential to ensure the persistent and ubiquitous activity of the genetic construct. Viral promoters such as the CMV promoter are often used for this purpose. Alternatively, a synthetic CAG promoter or house-keeping gene promoters such as Ubi and EF1a could be used.

Like ordinary genes, transgenes become targets for epigenetic systems regulating gene activity. Since constructs harbouring transgenes are usually assembled from several elements, often taken from different organisms (lentiviral backbone, CMV enhancer, etc.), it is difficult to predict how such an artificially assembled gene will interact with the cell's epigenetic systems. An undesirable outcome of such interaction is the silencing of the transgene which is a significant

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11248-022-00319-5>.

N. Battulin (✉) · A. Korablev · A. Ryzhkova · A. Smirnov · E. Kabirova · A. Khabarova · T. Lagunov · I. Serova · O. Serov
Laboratory of Developmental Genetics, Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia 630090
e-mail: battulin@gmail.com

N. Battulin
Institute of Genetic Technologies, Novosibirsk State University, Novosibirsk, Russia 630090

problem in animal studies, biotechnology and gene therapy (Alhaji et al. 2019).

Many promoter comparison studies in undifferentiated cells and diverse cell types were implemented to evaluate promoter strength and sensitivity to silencing (Qin et al. 2010), however there is much less data from animal experiments.

In this paper, we describe our experience with the EF1a promoter used to create transgenic mice for SARS-CoV-2 research. We reviewed several reports comparing different promoters and settled on the human EF1a promoter (Wang et al. 2008; Norrman et al. 2010; Zheng and Baum 2014). This constitutive promoter is widely recognized as one of the strongest expression drivers alongside CAG promoter (Qin et al. 2010; Norrman et al. 2010). According to tests on cell cultures, it provided a high constant expression in various cell types and applications. However, some transgenic animal studies demonstrate that the EF1a promoter might be subpar compared to CAG due to its susceptibility to silencing (Seita et al. 2019; Eun et al. 2020). We hope that our message will be helpful to specialists involved in creating transgenic animals.

Materials and methods

Transgenic mice generation

The EF1a-hACE2 construct was based on the pSBbi-GP plasmid. pSBbi-GP was a gift from Eric Kowarz (Addgene plasmid # 60,511) (Kowarz et al. 2015). First, the GFP-PuroR cassette was removed by restriction digest with PmlI and NdeI, end blunting and religation. The Ace2 gene (Addgene #1786) was then inserted at the NcoI cloning site with HiFi NEB cloning kit. Sleeping beauty mRNA was produced by *in vitro* transcription from the pCMV(CAT) T7-SB100 linearized vector (HiScribe™ T7 High Yield RNA Synthesis Kit, NEB) and mixed with the DrdI-linearized transposon vector (10 ng/μL Sleeping Beauty mRNA and 30 ng/μL DNA in TE buffer). pCMV(CAT)T7-SB100 was a gift from Zsuzsanna Izsvak (Addgene plasmid # 34,879) (Mátés et al. 2009). The solution was backfilled into an injection needle with positive balancing pressure (Transjector 5246, Eppendorf) and injected into the cytoplasm of zygotes (C57BL/6×CBA background). After

injections, the embryos were cultured for 1 h in drops of M16 medium at 37 °C and an atmosphere of 5% CO₂. The viable microinjected zygotes were transplanted the same day into oviducts of pseudopregnant CD-1 females (0.5 days after coitus). Isoflurane inhalation anesthesia was applied in these experiments.

Transgene insertion in mice was confirmed by PCR genotyping with primers for hACE2 gene and mouse SMC2 gene (Table 1). For ddPCR, primers and probes for hACE2 and mouse Emid1 or Usp17le reference genes (1 or 5 copies in mouse haploid genome, respectively) (Table 1) were used in accordance with manufacturer's protocol (ddPCR Supermix for Probes (No dUTP), BioRad). Droplet digital PCR (ddPCR) was performed using a QX100 system (BioRad). In brief, genomic DNA was digested overnight with MseI in CutSmart buffer (NEB) (1 μg genomic DNA in 30 μl) and added to the ddPCR mixture. ddPCR reactions were set in 20 μl volumes containing 1×ddPCR Supermix for Probes (no dUTP), 900 nM primers and 250 nM probes, and 1 μl of genomic DNA. Amount of DNA was based on the copy number for the exact line, and was in the range of 0.3–30 ng per reaction. ddPCR reactions for each sample were performed in duplicates. PCR was conducted according to the following program: 95 °C for 10 min, then 41 cycles of 95 °C for 30 s and 61 °C for 1 min, with a ramp rate of 2 °C per second, and a final step at 98 °C for 5 min. The results were analyzed using QuantaSoft software (Bio-Rad).

For RT-PCR, total RNA was extracted from mouse organs in glass homogenizers using TRI Reagent (Sigma-Aldrich). 2 μg of total RNA was used to generate cDNA in a 20 μl reaction using RevertAid RT Kit (Thermo Fisher Scientific) with random hexamer primers according to the manufacturer's instructions. RT-PCR was performed with the same primers used in ddPCR (see Table 1) for examination of the hACE2 inserts while the Rpl4 housekeeping gene was used as a control. 1 μl of resulting cDNA was used in a 20 μl RT-PCR reaction with 1×BioLabMix Taq buffer, 1.5 mM MgCl₂, 0.2 mM of each deoxynucleotide (dATP, dCTP, dGTP, and dTTP), 0.4 mM of forward and reverse primers, and 1 ul of HotStart Taq polymerase (BioLabMix, Russia). The reaction was conducted under the following conditions: initial denaturation 95 °C for 5 min, then 35 cycles of 95 °C for 15 s, 60 °C for 30 s, 72 °C for 30 s, and a final incubation at 72 °C for 2 min.

Table 1 Primer sequences for genotyping, ddPCR, RT-PCR, and bisulfite sequencing PCR of the EF-1 promoter region

Primer name	Sequence 5'–3'	PCR experiment
SeqEF1aprom-F	CCTTTTTGAGTTTGGATCTTGG	Genotyping (536 bp)
SeqACE2beg-R	GCATTCTTGTGGATTATCTGGG	
Rpl4-F	GGCTGCTTCCCTCAAGAGTA	RTPCR (121 bp)
Rpl4-R	AATCTTCTTGCCTGGTGCTC	
Ace2 F	TCCATGCTAACCGGACCCAGGAA	RTPCR and ddPCR (128 bp)
Ace2 R	TGAGCTGTCAGGAAGTCGTCCA	
Ace2 probe	HEX-CCACAGCTTGGGACCTGGGGAAGGG-BHQ2	
Rpl4 F	GGCTGCTTCCCTCAAGAGTA	RTPCR (121 bp)
Rpl4 R	AATCTTCTTGCCTGGTGCTC	
Emid1 F	GCCAGGACTGGGTAGCAC	ddPCR (79 bp)
Emid1 R	AGGAGGCTCCTGAATTTGTGACAAG	
Emid1 probe	FAM-CCTGGGTCATCTGAGCTGAGTCC-BHQ1	
Usp17le F	AGAACACAGGCAACAGCTGCT	ddPCR (111 bp)
Usp17le R	GGGAACAACAGGTTTGACTGT	
Usp17le probe	FAM-CCTGCAGTGCTTGACACACACACCACCTC-BHQ1	
BS_EF1A_F1	GTGTTAAGTTGGTTTAATTTTAGTTTTT	Bisulfite sequencing PCR
BS_EF1A_R1	CCTACTACAAAAAACTCAAATAAAAAA	
BS_EF1A_F2	TTTTATTTTGAGTTTTTTGTAGTAGGGT	Bisulfite sequencing PCR
BS_EF1A_R2	ACAAAATAATCTTATAAATACRAACCAA	

Animals were kept in a standard environment at 24 °C temperature, 40–50% relative air humidity and 14 h light/10 h dark–light-cycle. Food and water were available ad libitum. At the end of experiments, remaining animals were euthanized by CO₂. All experiments were conducted at the Centre for Genetic Resources of Laboratory Animals at the Institute of Cytology and Genetics, SB RAS (RFME-F161914X0005 and RFMEF161914X0010). All experiments were performed in accordance with protocols and guidelines approved by the Animal Care and Use Committee Federal Research Centre of the Institute of Cytology and Genetics, SB RAS operating under standards set by regulations documents Federal Health Ministry (2010/708n/RF), NRC and FELASA recommendations. Experimental protocols and euthanasia procedures were approved by the Bioethics Review Committee of the Institute of Cytology and Genetics. The manuscript followed the recommendations in the ARRIVE guidelines.

Cell culture and RBD staining

HEK293T cells were cultured at 37 °C under 5% CO₂ in DMEM (Thermo Fisher Scientific), supplemented

with 10% FBS (Capricorn Scientific), 1×penicillin & streptomycin 10x (Capricorn Scientific), 1×GlutaMax-I 100×(Thermo Fisher Scientific). For subculture cells were rinsed with 1×PBS and detached using 0.25% trypsin–EDTA at 37 °C for 3 min. Cells were typically split every 2–3 d at a 1:2 ratio.

For cell staining we used a biotinylated recombinant SARS-CoV-2 receptor-binding domain (RBD-bio) obtained by Taranin's group (Gorchakov et al. 2021). A total of ~0.5–1*10⁶ cells/ml were harvested using 0.25% trypsin–EDTA and washed in PBS. Cells were pellet by spinning at 1500×g for 10 min. Cell pellet was mixed with RBD-bio (2.55 mg/ml) at a final 0.5–1 µg per sample in an ice cold buffer (PBS, 10% FBS), and the cell suspension was incubated for at least 30 min on ice. After RBD-bio binding, the cell pellet was washed twice in an ice cold buffer, mixed with Avidin PE in an ice cold buffer in a ratio 1:600 and incubated for 10 min on ice. The pellet was washed twice in an ice cold buffer to minimize unnecessary staining. For a positive control, HEK293T cells expressing SARS-CoV-2 S protein on their surface from Taranin's group were used. For a negative control, cell suspension was washed in PBS without further staining. The cell staining was

visualized using ZOE fluorescent cell imager (Bio-Rad) or run on a FACS Aria III (Becton Dickinson).

mES cells were grown at 37 °C in a humidified atmosphere of 5% CO₂ and passaged on average every 2–3 days. For passaging, cells were dissociated with 0.05% Trypsin–EDTA and gentle trituration. mESCs were cultured on plates coated with 0.1% gelatin under 2i condition (1 μM PD, 3 μM CHIR) in DMEM (ThermoFisher), supplemented with 7.5% ES FBS (Gibco), 7.5% KSR (Gibco), 1 mM L-glutamine (Sigma), NEAA (Gibco), 0.1 mM β-mercaptoethanol, LIF (1000 U/ml, Polygen) and Pen-Strep (100 U ml⁻¹ each).

We have previously generated mESCs with randomly inserted EF1-AtAFB2-Cherry-IRES-PuroR construct (Yunusova et al. 2021). Integrated transgenes underwent partial silencing with passaging in culture, resulting in mosaic expression in one-cell derived mESCs colonies, which contained both Cherry⁺ and Cherry⁻ cells. Cherry⁺ and Cherry⁻ populations were sorted using a FACS Aria III (Becton Dickinson).

DNA isolation and bisulfite sequencing PCR amplification

DNA from cultured cells and animal tissue was isolated by digestion with proteinase K and phenol–chloroform extraction for further bisulfite conversion. Approximately 1 μg of DNA was used for bisulfite modification using EZ DNA Methylation-Gold kit (Zymo Research) following the manufacturer's protocol.

EF-1 promoter regions were amplified using primer pairs designed using The Bisulfite Primer Seeker Zymo Research software (Table 1). During the first round of PCR, 1 μl of bisulfite treated DNA template was amplified using the HS-Taq PCR kit (Bio-labmix). PCR products (1 μl) were then subjected to a second round of amplification (20 μl total reaction mixture) to obtain an appropriate amount of DNA for further sequencing. PCR products were visualized on agarose gel. The relative positions of primers and CpG dinucleotides within the promoter sequence is shown in Fig. 1A. PCR products were purified with Ampure XP beads (Beckman Coulter) and samples were prepared for sequencing using KAPA Hyper-Prep Kit (KAPA Biosystems).

Bisulfite sequencing data analysis

Sequencing data was aligned to the in silico bisulfite converted human genome assembly hg38 using Bismark aligner v0.23 (Krueger and Andrews 2011) with default settings (bismark –path_to_bowtie2 path_to_bowtie2 –local -n 1 –genome path_to_genome –non-directional -B file.fastq.gz). Reads were aligned in single end modus with each read of a paired-end read aligned separately.

Alignments were merged together and duplicates were removed from aligned reads using deduplicate_bismark. After deduplicating, alignments were filtered by MAPQ 10.

Methylation status of all cytosines was extracted using bismark_methylation_extractor (bismark_methylation_extractor –report –gzip –multicore 4 –bedGraph –CX file.deduplicated.bam). Coverage of cytosines was extracted using coverage2cytosine (coverage2cytosine –genome_folder path_to_genome -o output_name file.deduplicated.bismark.cov.gz). BedGraph files with positional methylation were made from {name}.CpG_report.txt file by dividing the 4th column by the sum of 4th and 5th (4th—count methylated; 5th—count non-methylated).

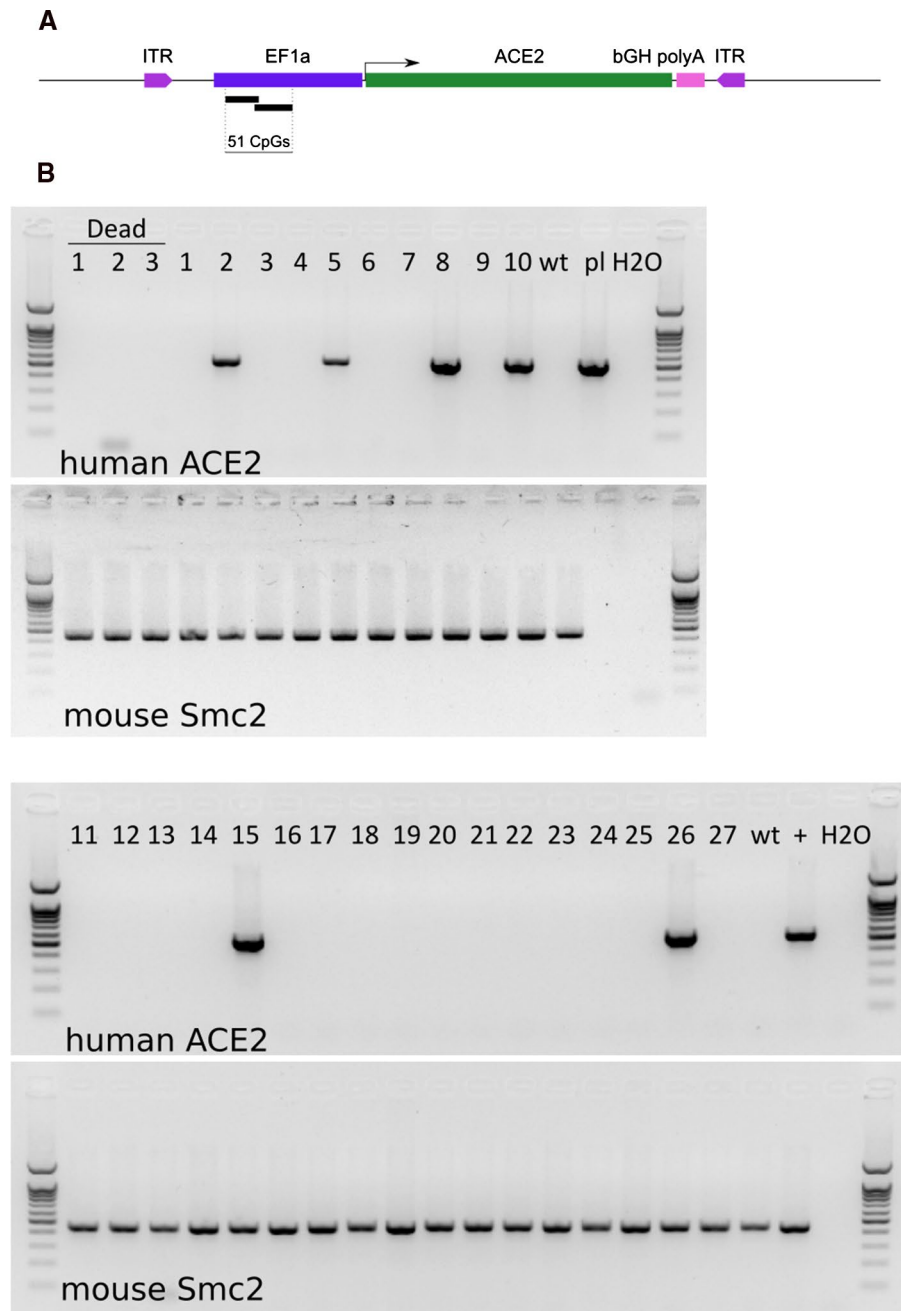
Results

Obtaining transgenic animals

Concatemer is a typical feature of conventional pronuclear microinjection. To obtain single copy inserts in addition to the concatemers, we used the Sleeping Beauty transposon system to deliver our genetic construct in a genome (Ivics et al. 2014). Figure 1A shows the general design of the genetic construct used in our experiment (see also Sup. Figure 1).

For microinjection, we linearized the genetic construct. Using a linear construct in the Sleeping Beauty technology increases the transgenesis efficiency up to 70% (Dr. Boris V Skryabin personal communication). During microinjection into mouse zygotes obtained by in vitro fertilization, a mix containing transposon DNA and transposase mRNA was injected into the cytoplasm. After embryo transfer to pseudopregnant females, we obtained 31 pups, of which six, according to PCR genotyping, carried the transgene integration

Fig. 1 Generation of transgenic mice. **A** Scheme of the elements of the transgenic construct. ITR – inverted repeats of Sleeping Beauty transposon; EF1a – promoter; ACE2 – cDNA of human ACE2; bGH polyA – bovine growth hormone polyadenylation signal; two black bars indicate the positions of amplicons for bisulfite sequencing, the data for which are presented in Fig. 3. **B** Results of PCR genotyping of mice born after pronuclear microinjection. Primers for hACE2 were used for detection of the transgene in the genome; mouse Smc2—internal DNA loading control; wt – genomic DNA from non-transgenic mice; + genomic DNA from transgenic mice; pl – plasmid with ACE2 cDNA



(Fig. 1B). Thus, the efficiency of obtaining transgenic animals was about 20% of those born.

Transgenic founders were crossed with C57BL/6 animals to generate F1 pups for analysis. Two founders, #5 and 10, appeared to be infertile, while another four (#2, 8, 15, 26) successfully transmitted the transgene (Sup. Figure 2). Transgene copy number analysis for founders by ddPCR revealed that in most

cases animals have a multiple copies of transgenes with >60 copies, except #2 where integrated transgene had three copies (Sup. Figure 2, 3). Such a large number of transgene copies per genome can be explained by assuming that, along with single integrations of transposons into the genome, insertion of a multicopy concatemer, typical of pronuclear microinjection, could occur (Smirnov and Battulin 2021).

Although ddPCR could be imprecise when applied to cases with > 100 copy integrations, we at least could confirm separate integration sites (Sup. Figure 2, 3). For instance, in animals from line #8, there were two independent sites with ~60 and ~150 copies. In line #15 there were at least three integration sites, including one with > 1000 copies. Independent one copy integrations were also detected in two lines (#15, 26).

Evaluation of transgene activity

To qualitatively assess the expression of the hACE2 gene in the tissues of four F1 transgenic animals with independent integrations events (Sup. Figure 2, green highlight), we isolated RNA from several organs (spleen, lung, heart, liver) and performed RT-PCR (Fig. 2A). We found that only animals from

the F0#15 founder's offspring showed some low transcription levels of the hACE2 gene. For all other animals analyzed, RT-PCR could not detect the presence of human gene transcripts in tissues.

We also evaluated the hACE2 protein levels. In order to visualize the presence of the protein on the cell surface, we used the method of cell staining based on a recombinant coronavirus receptor-binding (RBD) S protein conjugated with biotin and detection with a fluorescent-labeled avidin (Gorchakov et al. 2021). First, we validated that the EF1a-hACE2 construct is capable of providing a high level of production of the hACE2 protein. We transfected HEK293 cells with the EF1a-hACE2 plasmid (Fig. 1A) and stained them with RBD dye, discovering many hACE2-positive cells (Fig. 2B). Although expression from transfected plasmid DNA cannot be compared

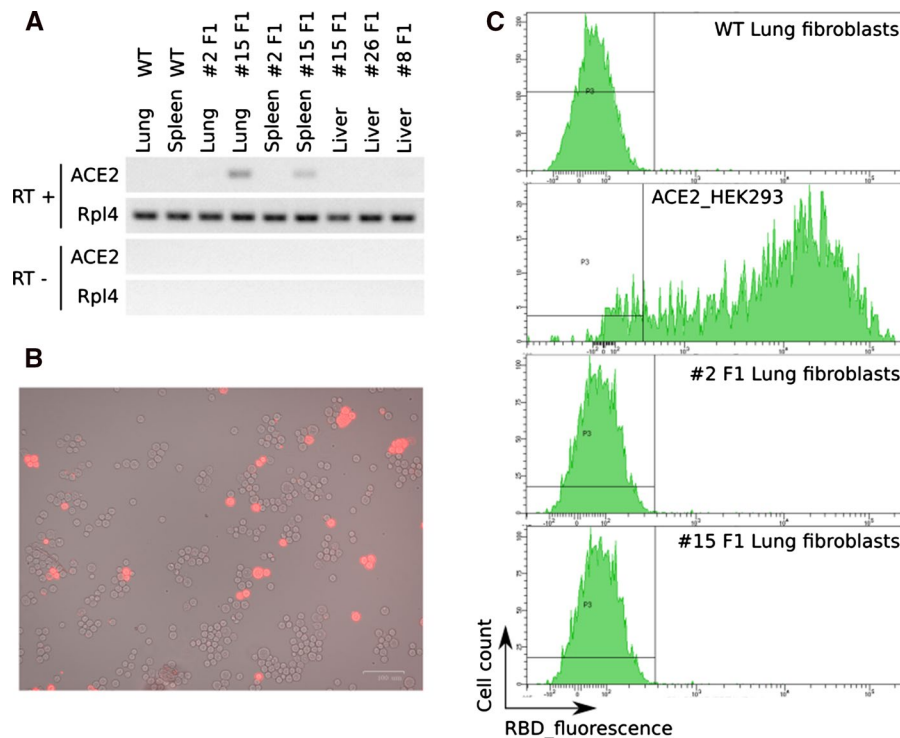


Fig. 2 Activity of the hACE2 transgene in transgenic animals. **A** RT-PCR analysis of hACE2 in organs derived from heterozygous F1 animals. Rpl4 – housekeeping gene for internal control; RT- samples without RT for DNA contamination control; WT- non-transgenic animal; #2 F1—heterozygous F1 animal from founder #2. **B** Staining of HEK293 cells with recombinant biotin conjugated RBD of SARS-CoV-2 and a fluorescent-labeled avidin. HEK293 cells were previously transfected with the EF1a-hACE2 plasmid, which was used for

microinjection in obtaining transgenic animals. Not all cells are stained because only a portion of cells receive the plasmid upon transfection. **C** FACS analysis of lung fibroblasts stained with recombinant biotin conjugated RBD of SARS-CoV-2 and a fluorescent-labeled avidin. WT lung fibroblasts used as negative control, HEK293 cells transfected with the EF1a-hACE2 plasmid as a positive control. No positive cells were detected in fibroblast cultures derived from heterozygous F1 animals

directly to genomic transgene insertion, this experiment verified that the designed construct is indeed functional and facilitates hACE2 expression.

Likewise, RBD staining was performed to look for the hACE2 on the cell surface of primary cultures of lung fibroblasts and peritoneal cells derived from F1 animals from the six transgenic lines. Unfortunately, none of the animals showed the presence of hACE2, including the animal (from the founder F0#15) with RT-PCR signal (Fig. 2C).

We hypothesized that the lack of expression from the transgene might be due to the epigenetic silencing of the promoter. It is known that promoter DNA methylation is often associated with silencing. Therefore, we assessed promoter methylation level using the bisulfite sequencing method. We isolated DNA from the tail tips of transgenic animals and determined the level of methylation for 51 CpG dinucleotides in the promoter (their localization in the construct is indicated in Fig. 1). We discovered that the average promoter methylation level among all animals was about 90% (Fig. 3). To test whether the DNA methylation level of the EF1a promoter is indeed associated with a change in transcription activity, we used an embryonic stem cells (ESCs) line available in the laboratory, with a fluorescent mCherry protein expressed under the EF1a control (Yunusova et al. 2021). It is a well-established fact that integrated genetic constructs could lose expression in the absence of selective pressure (Martin and Whitelaw 1996; Whitelaw et al. 2001). The selected ES cell line is characterized by the spontaneous switch-off of the construct

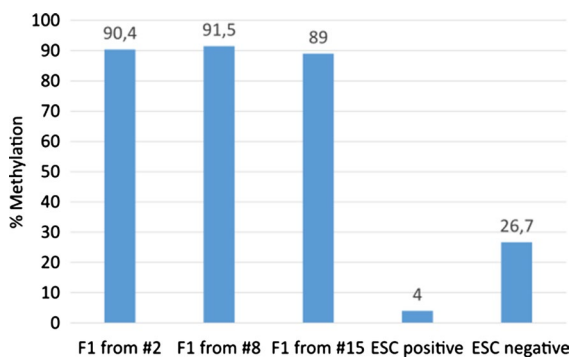


Fig. 3 DNA methylation levels of the EF1a promoter in transgenic animals. The analyzed region is indicated in Fig. 1a. ESC positive/negative – cell populations with active or silent EF1a promoter (Cherry[±]), respectively

so that the culture constantly contains subpopulations of cells with or without transgene expression (mCherry[±]). We sorted two mCherry subpopulations and examined the level of EF1a promoter methylation. In the cell population with the active construct, methylation level was only 4%. In cells with silenced transcription, this number increased to about 25%. This experiment suggests that the DNA methylation level of the EF1a promoter correlates with its activity and that levels as much as 25% are sufficient for silencing.

Discussion

Proper choice of the promoter is important for obtaining transgenic animals with robust expression. The nominal rating of the constitutive mammalian promoter has been suggested to be CAG>EF1a>CMV>SV40>PGK>UbC (Qin et al. 2010; Seita et al. 2019). For our experiments, we chose human elongation factor (EF1a)—one of the most popular promoters. Cell culture experiments generally favor EF1a promoter for human and mouse expression (Kim et al. 2007; Norrman et al. 2010). The promoter sequence used in the study was initially cloned in the bidirectional vector, pSBbi-GFP, and demonstrated high activity in many cell studies (Kowarz et al. 2015; Widera et al. 2021; Yeung et al. 2021). Prior to experiments we removed the GFP-Puro cassette to avoid transcription interference and used EF1a-hACE2 fragment for microinjections (Fig. 1A).

Unfortunately, in our case, transgenic animals demonstrated only a very low expression level in one of four tested lines. Theoretically, random integration of transgenes could lead to inactivation due to position effect variegation (integration in the inactive locus) or repeat-induced gene silencing (Garrick et al. 1998; Smirnov and Battulin 2021). Analysis of multiple founders usually helps to alleviate these problems and find animals with acceptable expression levels. We screened F1 animals from four founders (Sup. Figure 2) and did not find high expression cases. This might indicate a global regulation problem. One of the possible culprits in such cases is epigenetic silencing (Garrick et al. 1998; Matzke et al. 2000), when the introduced genetic construct is switched off due to DNA hypermethylation and decoration with repressive histone marks (Jones 2012; Moore et al.

2013). It is most likely that silencing occurs in early embryonic development during global epigenetic reprogramming (Du et al. 2021). Therefore, a positive test result of a genetic construct on a cell culture (for example, fibroblasts) does not guarantee that this construct will also be active in this cell type (fibroblasts) in the body of a transgenic animal. It is possible that some sequences in the construct will undergo epigenetic silencing in early development, and this developmental background will not allow the construct to be transcribed when differentiated into the target cell type. This notion is common for CMV promoter, as some experiments in animals and cells show it could be severely silenced compared to CAG promoter (Xia et al. 2007; Duan et al. 2012). As far as we know, it is impossible to predict in advance whether elements of the transgenic construct will be subjected to silencing or not, therefore this information can only be obtained in an experiment on animals.

Another element of the transgene potentially subject to silencing is the terminal repeats of the Sleeping Beauty transposon. However, there is evidence against this possibility. Garrels et al. analyzed 67 independent integrations of the Sleeping Beauty transposon with the CAG-Venus construct in transgenic mice and showed that 66 integrations resulted in ubiquitous promoter-dependent expression of the reporter (Garrels et al. 2016). Thus, this large sample confirms that the Sleeping Beauty transposon can provide stable expression of the constitutive promoter.

We found this recommendation from Brian Sauer's lab in a plasmid comment on the Addgene website (<https://www.addgene.org/11918>): "However, pBS513 is not suitable for use in transgenic mice as such EF-1 alpha transgene constructs are often inactive in many tissues". We believe that this information should be explicitly published in a peer-reviewed journal, as it may help researchers avoid problems when creating transgenic animals.

Literature analysis reveals that EF1a promoter's performance in transgenic animals is controversial. Although there are many successful studies, sometimes promoter silencing is observed. Some notable cases are discussed below.

It is pertinent to note that the first work in which the EF1a promoter was used to drive expression in transgenic mice emphasized its strength and ubiquity of expression (Hanaoka et al. 1991). These conclusions were well confirmed by experiments.

However, the authors selected the founders for the transgenic line based on the activity of the chloramphenicol acetyltransferase (CAT) enzyme, and not on the presence of the transgene in the genome. Of the 102 born after microinjection of the pEF321 CAT construct, the authors found three with CAT activity in the tail, one of which became the founder of a line on which the activity of the transgene in organs was tested. Since that work did not screen for the presence of the transgene in the genome, we can postulate that survivorship bias occurred in this case. This means that the authors concluded that the promoter activity was ubiquitous by selecting a founder in which the insert was active, while founders with an inactive transgene were probably not included in the study. Note that this is only a hypothesis that we cannot confirm.

Chevalier-Mariette et al. have evaluated how the transgene expression activity is influenced by its CpG content (Chevalier-Mariette et al. 2003). Two reporter genes, differing in their CpG content, were constructed from a CpG-rich *LacZ* gene—*LagZ* and *LagoZ*. To obtain these transgenes the authors carried out codon optimization, which reduced the CpG content of the transcribed region from 9.24% in *LacZ* to 1.6% in *LagZ* and 0.06% in *LagoZ*. EF1a promoter was used in all constructs. The differences in expression patterns between these transgenes were dramatic. No activity of the *EF1aLacZ* transgene was observed in somatic tissues from the nine independent mouse lines. However, a widespread expression of CpG-null *LagoZ* transgene under the control of the same EF1a promoter was observed in four independent mouse lines. These results indicate that the readout of EF1a promoter activity can be dependent on the CpG density of the transgene. However the transcribed region of *ACE2* in our construct has a CpG content of 1.1%. Therefore, expression silencing is probably not related to the CpG density of the transcribed region in our case, although we cannot completely rule out this possibility.

Interestingly, a unique case of epigenetic regulation similar to the regulation of imprinted genes has been described for a transgenic construct with the EF1a promoter (Uchiyama et al. 2014). A line of transgenic animals EF1a-LacZ was found with the transgene active only in case of paternal inheritance. This effect is explained by the influence of the genome elements neighboring the transgene integration site.

The EF1a promoter has been used to drive reporter gene expression in several studies in which the transgene was delivered into the genome with lentivirus vectors (Sehgal et al. 2011; Li et al. 2013; Qin et al. 2015). However, Qin et al. noted that despite the presence of the EF1a-GFP transgene in the genome, expression was not detected either by RT-PCR or by fluorescence microscopy (Qin et al. 2015). In the work of Li et al., transcription was detected only in the testis and no other organs (Li et al. 2013). In contrast to these works, Sehgal et al. showed expression of the EF1a-GFP construct in multiple organs in at least one transgenic mice line (Sehgal et al. 2011).

Eun et al. used the human EF1a promoter to generate a dog with ubiquitous GFP expression (Eun et al. 2020). The authors introduced the EF1a-GFP transgene construct into fibroblasts and obtained a cell culture that stably expresses GFP. These cells were used in a somatic cell nuclear transfer procedure to obtain cloned animals with EF1a-GFP. However, analysis of four born transgenic puppies did not reveal GFP expression at the organismal and cellular levels in these transgenic dogs. Interestingly, treatment of the fibroblast culture derived from a cloned animal with a DNA methyltransferase inhibitor 5-Azacytidine, and a histone deacetylase inhibitor, Trichostatin A resulted in reactivation of the transgene. This experiment confirms that constructs containing the EF1a promoter can undergo epigenetic silencing in embryonic development.

Side-by-side comparison of the two GFP constructs with CAG and EF1a promoters was performed in the transgenic monkey project (Seita et al. 2019). Authors took a stringent approach to characterize the expression of the GFP transgene at multiple developmental stages: from cultured ES cells and blastocysts to specialized tissues. They concluded that although both promoters were active in undifferentiated ES cells, the EF1a promoter was gradually silenced during differentiation (Seita et al. 2019). In one of the transgenic monkey lines with four copies the EF1-GFP construct was inactive, which was attributed to local genomic effects.

Thus, our data, as well as some published results, suggest that EF1a promoter can indeed become silent in development, so using it as a strong constitutive ubiquitous promoter may not be the best choice for transgenic animal models.

Conclusions

Obtaining transgenic mice for research can be a challenging task, as demonstrated in this study, and promoter choice is a critical component for a project's success. However, selection of the promoter based on the literature data could be difficult and requires in-house tests in the lab. This is partly caused by the publication bias, because negative results are less likely to be published. Ideally, there must be a side-by-side comparison of popular promoters in transgenic animals with expression monitoring for a few generations. We think that such efforts combined with extensive meta study of the published cases would help to create a set of recommendations for selecting an optimal promoter, which will help to streamline animal experiments.

Institutional Review Board Statement

The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Bioethical Committee at the Institute of Cytology and Genetics (protocol N45 11.16. 2018).

Author's Contribution Conceptualization, N.B., A.K. and O.S.; methodology, A.K., A.R., A.S., E.K., A.Kh., T.L., I.S.; writing—original draft preparation, N.B.; writing—review and editing, A.S. and N.B.; supervision, N.B. and O.S.; funding acquisition, N.B. and O.S. All authors have read and agreed to the published version of the manuscript.

Funding This work was supported by Russian Foundation for Basic Research, grant number 20-04-60094. N.B. supported from the Ministry of Education and Science of Russian Federation, grant #2019-0546 (FSUS-2020-0040). Data analysis performed on computational nodes of the Institute of Cytology and Genetics (state project FWN-2022-0019).

References

- Alhaji SY, Ngai SC, Abdullah S (2019) Silencing of transgene expression in mammalian cells by DNA methylation and histone modifications in gene therapy perspective. *Bio-technol Genet Eng Rev*. <https://doi.org/10.1080/02648725.2018.1551594>
- Chevalier-Mariette C, Henry I, Montfort L et al (2003) CpG content affects gene silencing in mice: evidence from novel transgenes. *Genome Biol* 5:83

- Du Z, Zhang K, Xie W (2021) Epigenetic reprogramming in early animal development. *Cold Spring Harb Perspect Biol.* <https://doi.org/10.1101/cshperspect.a039677>
- Duan B, Cheng L, Gao Y et al (2012) Silencing of fat-1 transgene expression in sheep may result from hypermethylation of its driven cytomegalovirus (CMV) promoter. *Theriogenology* 78:793–802. <https://doi.org/10.1016/j.theriogenology.2012.03.027>
- Eun K, Hong N, Jeong YW et al (2020) Transcriptional activities of human elongation factor-1 α and cytomegalovirus promoter in transgenic dogs generated by somatic cell nuclear transfer. *PLoS ONE.* <https://doi.org/10.1371/journal.pone.0233784>
- Garrels W, Talluri TR, Ziegler M et al (2016) Cytoplasmic injection of murine zygotes with sleeping beauty transposon plasmids and minicircles results in the efficient generation of germline transgenic mice. *Biotechnol J* 11:178–184. <https://doi.org/10.1002/biot.201500218>
- Garrick D, Fiering S, Martin DIK, Whitelaw E (1998) Repeat-induced gene silencing in mammals. *Nat Genet* 18:56–59. <https://doi.org/10.1038/ng0198-56>
- Gorchakov AA, Kulemzin SV, Guselnikov SV et al (2021) Isolation of a panel of ultra-potent human antibodies neutralizing SARS-CoV-2 and viral variants of concern. *Cell Discov* 7:96. <https://doi.org/10.1038/s41421-021-00340-8>
- Hanaoka K, Hayasaka M, Uetsuki T et al (1991) A stable cellular marker for the analysis of mouse chimeras: the bacterial chloramphenicol acetyltransferase gene driven by the human elongation factor 1 α promoter. *Differentiation.* <https://doi.org/10.1111/j.1432-0436.1991.tb00256.x>
- Ivics Z, Mátés L, Yau TY et al (2014) Germline transgenesis in rodents by pronuclear microinjection of sleeping beauty transposons. *Nat Protoc.* <https://doi.org/10.1038/nprot.2014.008>
- Jones PA (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 13:484–492. <https://doi.org/10.1038/nrg3230>
- Kim S, Gi JK, Miyoshi H et al (2007) Efficiency of the elongation factor-1 α promoter in mammalian embryonic stem cells using lentiviral gene delivery systems. *Stem Cells Dev.* <https://doi.org/10.1089/scd.2006.0088>
- Kowarz E, Löscher D, Marschalek R (2015) Optimized Sleeping Beauty transposons rapidly generate stable transgenic cell lines. *Biotechnol J* 10:647–653. <https://doi.org/10.1002/biot.201400821>
- Krueger F, Andrews SR (2011) Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 27:1571–1572. <https://doi.org/10.1093/bioinformatics/btr167>
- Li X, Mao Z, Wu M, Xia J (2013) Rescuing infertility of pick1 knockout mice by generating testis-specific transgenic mice via testicular infection. *Sci Rep.* <https://doi.org/10.1038/srep02842>
- Martin DIK, Whitelaw E (1996) The vagaries of variegating transgenes. *BioEssays* 18:919
- Mátés L, Chuah MKL, Belay E et al (2009) Molecular evolution of a novel hyperactive sleeping beauty transposase enables robust stable gene transfer in vertebrates. *Nat Genet* 41:753–761. <https://doi.org/10.1038/ng.343>
- Matzke MA, Mette MF, Matzke AJM (2000) Transgene silencing by the host genome defense: implications for the evolution of epigenetic control mechanisms in plants and vertebrates. *Plant Mol Biol.* <https://doi.org/10.1023/a:1006484806925>
- Moore LD, Le T, Fan G (2013) DNA Methylation and Its Basic Function. *Neuropsychopharmacology* 38:23–38. <https://doi.org/10.1038/npp.2012.112>
- Norrman K, Fischer Y, Bonnamy B et al (2010) Quantitative comparison of constitutive promoters in human ES cells. *PLoS ONE* 5:e12413. <https://doi.org/10.1371/journal.pone.0012413>
- Qin J, Xu H, Zhang P et al (2015) An efficient strategy for generation of transgenic mice by lentiviral transduction of male germline stem cells in vivo. *J Anim Sci Biotechnol.* <https://doi.org/10.1186/s40104-015-0058-4>
- Qin JY, Zhang L, Clift KL et al (2010) Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. *PLoS ONE* 5:e10611. <https://doi.org/10.1371/journal.pone.0010611>
- Sehgal L, Thorat R, Khapare N et al (2011) Lentiviral mediated transgenesis by in vivo manipulation of spermatogonial stem cells. *PLoS ONE.* <https://doi.org/10.1371/journal.pone.0021975>
- Seita Y, Tsukiyama T, Azami T et al (2019) Comprehensive evaluation of ubiquitous promoters suitable for the generation of transgenic cynomolgus monkeys†. *Biol Reprod* 100:1440–1452. <https://doi.org/10.1093/biolre/iox040>
- Smirnov A, Battulin N (2021) Concatenation of transgenic DNA: random or orchestrated? *Genes (basel)* 12:1969. <https://doi.org/10.3390/genes12121969>
- Uchiyama K, Watanabe D, Hayasaka M, Hanaoka K (2014) A novel imprinted transgene located near a repetitive element that exhibits allelic imbalance in DNA methylation during early development. *Dev Growth Differ.* <https://doi.org/10.1111/dgd.12182>
- Wang R, Liang J, Jiang H et al (2008) Promoter-dependent EGFP expression during embryonic stem cell propagation and Differentiation. *Stem Cells Dev* 17:279–290. <https://doi.org/10.1089/scd.2007.0084>
- Whitelaw E, Sutherland H, Kearns M et al (2001) Epigenetic effects on transgene expression gene knockout protocols. Humana Press
- Widera M, Wilhelm A, Toptan T et al (2021) Generation of a sleeping beauty transposon-based cellular system for rapid and sensitive screening for compounds and cellular factors limiting SARS-CoV-2 replication. *Front Microbiol.* <https://doi.org/10.3389/fmicb.2021.701198>
- Xia X, Zhang Y, Zieth CR, Zhang S-C (2007) Transgenes delivered by lentiviral vector are suppressed in human embryonic stem cells in a promoter-dependent manner. *Stem Cells Dev* 16:167–176. <https://doi.org/10.1089/scd.2006.0057>
- Yeung TK, Lau HW, Ma HT, Poon RYC (2021) One-step multiplex toolkit for efficient generation of conditional gene silencing human cell lines. *Mol Biol Cell.* <https://doi.org/10.1091/mbc.E21-02-0051>
- Yunusova A, Smirnov A, Shnaider T et al (2021) Evaluation of the OsTIR1 and AtAFB2 AID systems for genome

architectural protein degradation in mammalian cells. *Front Mol Biosci.* <https://doi.org/10.3389/fmolb.2021.757394>

Zheng C, Baum BJ (2014) All human EF1 α promoters are not equal: markedly affect gene expression in constructs from different sources. *Int J Med Sci* 11:404–408. <https://doi.org/10.7150/ijms.8033>

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.