Comprehensive evaluation of ubiquitous promoters suitable for the generation of transgenic cynomolgus monkeys

Yasunari Seita[1], Tomoyuki Tsukiyama[1, 2], Takuya Azami[1], Kenichi Kobayashi[1,3], Chizuru Iwatani[1], Hideaki Tsuchiya[1], Masataka Nakaya[1], Hideyuki Tanabe[4], Seiji Hitoshi[5], Hiroyuki Miyoshi[6], Shinichiro Nakamura[1], Akihiro Kawauchi[3], and Masatsugu Ema[1, 2, 7*]

Short title: Promoters for ubiquitous Tg expression in NHPs

Summary Sentence: Comparative examination of ubiquitous promoters in an *in vitro* ES cell differentiation system and Tg animals provides fundamental insights into suitable promoters for high level transgene expression in cynomolgus monkeys.

[1] Department of Stem Cells and Human Disease Models, Research Center for Animal Life Science, Shiga University of Medical Science, Seta, Tsukinowa-cho, Otsu, Shiga 520-2192, Japan

[2] Institute for the Advanced Study of Human Biology (ASHBi), Kyoto University, 606-8501, Japan

[3] Department of Urology, Shiga University of Medical Science, Seta, Tsukinowa-cho, Otsu, Shiga 520-2192, Japan

[4] Department of Evolutionary Studies of Biosystems, School of Advanced Sciences, SOKENDAI (The Graduate University for Advanced Studies), Shonan Village, Hayama, Kanagawa 240-0193, Japan

[5] Department of Physiology, Shiga University of Medical Science, Seta, Tsukinowa-cho, Otsu, Shiga 520-2192, Japan

[6] Department of Physiology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan

[7] PRESTO, Japan Science and Technology Agency, Saitama, 332-0012, Japan

[*]To whom correspondence should be addressed.

Department of Stem Cells and Human Disease Models, Research Center for Animal Life Science, Shiga University of Medical Science, Seta, Tsukinowa-cho, Otsu, Shiga 520-2192, Japan. Phone: +81-77-548-2334; Fax: +81-77-543-1990;

E-mail: mema@belle.shiga-med.ac.jp

- **Grant support:** This study was supported in part by JSPS KAKENHI Grant Numbers JP17K14977 and JP17K08135 to Y.S. and T.T., respectively, PRESTO, Japan Science and Technology Agency to M.E., and a grant from Shiga University of Medical Science to M.E., Y.S., and T.T.

- **Keywords:** Blastocyst, Embryonic stem cells, Differentiation, Intracytoplasmic sperm injection (ICSI), Primates, Zygote

Abstract

Nonhuman primates (NHPs) are considered to be the most valuable models for human transgenic (Tg) research into disease, because human pathology is more closely recapitulated in NHPs than rodents. Previous studies have reported the generation of Tg NHPs that ubiquitously overexpress a transgene using various promoters, but it is not yet clear which promoter is most suitable for the generation of NHPs overexpressing a transgene ubiquitously and persistently in various tissues. To clarify this issue, we evaluated four putative ubiquitous promoters, cytomegalovirus (CMV) immediate-early enhancer and chicken beta-actin (CAG), Elongation factor 1α (EF1α), Ubiquitin C (UbC), and CMV, using an *in vitro* differentiation system of cynomolgus monkey

embryonic stem cells (ESCs). While the EF1α promoter drove Tg expression more strongly than the other promoters in undifferentiated pluripotent ESCs, the CAG promoter was more effective in differentiated cells such as embryoid bodies and ESC-derived neurons. When the CAG and EF1α promoters were used to generate green fluorescent protein (GFP)-expressing Tg monkeys, the CAG promoter drove GFP expression in skin and hematopoietic tissues more strongly than in EF1α-GFP Tg monkeys. Notably, the EF1α promoter underwent more silencing in both ESCs and Tg monkeys. Thus, the CAG promoter appears to be the most suitable for ubiquitous and stable expression of transgenes in the differentiated tissues of Tg cynomolgus monkeys and appropriate for the establishment of human disease models.

Introduction

Many animal models have been created to better understand human diseases and to develop novel therapies. Rodents have been used as model organisms for decades, because they share common anatomical and pathophysiological features with humans in the nervous, cardiovascular, and other internal organ systems [1], and rodent disease models are relatively easy to be generated by pronuclear microinjection of a transgene [2], by genome-engineered embryonic stem cells (ESCs) [3,4] and by genome editing in fertilized oocytes [5,6]. However, rodents do not always recapitulate human disease conditions. For example, although humans suffering from Parkinson Disease (PD) show spontaneous degeneration of dopaminergic neurons in the substantia nigra and behavioral abnormalities, mice lacking a causative gene of familial PD such as *Parkin, Pink1* [7], *Dj-1* [8] and *Lrrk2* [9] show no abnormalities in dopamine-related behavior [7,8,10,11]. Even triple gene knockout mice lacking *Parkin*, *DJ-1* and *Pink1* have normal morphology and numbers of dopaminergic and noradrenergic neurons in the substantia

nigra [12]. Moreover, transcriptomes exhibit substantial differences between mice and humans; thus, more than 20% protein-coding genes are not orthologous (reviewed by [13]). Another study clearly showed that the genetic and cellular mechanisms of human germ cell development are substantially different from rodents [14]. Therefore, animal models to recapitulate human development and disease phenotypes more faithfully are needed [15].

Nonhuman primates (NHPs) are considered to be the most valuable models for human diseases, because NHPs are closer to humans in terms of physiology [16], neurology [17] and genetics [18] than rodents. NHPs comprise the New World and Old World monkeys. The former includes common marmosets (*Callithrix jacchus)*; the latter includes macaques such as cynomolgus monkeys (*Macaca fascicularis*), rhesus monkeys (*Macaca mulatta*), and Japanese monkeys (*Macaca fuscata*). Macaques are closer genetically to humans than marmosets [19]. To create NHP models of human diseases, transgenic (Tg) monkeys have been created by using various promoters in viral-based vectors [20–24]. When the elongation factor 1 alpha (EF1α) promoter was used to introduce the gene encoding green fluorescent protein (GFP) into rhesus monkeys, GFP expression was detected only in the placenta and umbilical cord, but it was not detected in fetal tissues [20,23]. Niu *et al*. used the human ubiquitin C (UbC) promoter to overexpress GFP and obtained Tg rhesus monkeys expressing GFP in the skin successfully [21]. Sasaki *et al*. tried to create Tg marmosets with a cytomegalovirus (CMV) immediate-early enhancer and chicken beta-actin (CAG) promoter, a CMV promoter and EF1α promoter, and obtained CAG-GFP-Tg and CMV-GFP-Tg animals overexpressing GFP in blastocysts and various tissues [22]. Tomioka et al established a tetracycline-induced transgene expression system in marmosets [24]. However, it is not

clear which promoter is most suitable for persistent and ubiquitous overexpression in undifferentiated and differentiated tissues of Tg NHPs.

Here, we aimed to clarify promoters suitable for ubiquitous and stable Tg overexpression in the tissues of cynomolgus monkeys. We evaluated four ubiquitously used promoters—CAG, EF1α, UbC and CMV—by combining cynomolgus monkey embryonic stem cells (ESCs) with targeted knock-in of a promoter into the Adeno-Associated Virus Integration Site 1 (AAVS1) locus and using an *in vitro* ESC differentiation system. Whereas the EF1α promoter drove transgenic expression more strongly than other promoters in undifferentiated pluripotent ESCs, the CAG promoter was more effective in differentiated cells such as embryoid bodies and ESC-derived neurons. The CAG promoter drove GFP expression more strongly than the EF1α promoter in the skin and hematopoietic tissues of Tg animals, but less strongly than the EF1α promoter in blastocysts. The EF1α promoter was also susceptible to epigenetic silencing in both ES cells and animals. These data show that the CAG promoter is most suitable for the ubiquitous and stable expression of transgenes in undifferentiated and differentiated tissues of Tg cynomolgus monkeys.

Materials and Methods

Construction of targeting vector for ESCs

Plasmids expressing humanized CRISPR-associated protein 9 (hCas9) and single guide (sg)RNA (5′-CAC CTA TAA GGT GGT CCC GGC TCT-3′, 5′-CAC CGA GGA CCG ATT AAT ATG GCT C-3′) were prepared by ligating oligos into the BbsI restriction enzyme site of pX330 (http://www.addgene.org/42230/) [25]. The donors (pBR-monkeyAAVS1-FRT-SA-IZpA-FRT-Rev_CAG_EGFP_pA-MC1DTApAII, pBR-monkeyAAVS1FRT-SA-IZpA-FRT-Rev_EF1α_EGFP_pA-MC1DTApAII, pBR-monkeyAAVS1-FRT-SA-IZpA-FRT-Rev_UbC_EGFP_pA-MC1DTApAII and pBR-monkeyAAVS1-FRT-SA-IZpA-FRT-Rev_CMV_EGFP_pA-MC1DTApAII) were constructed by introducing the left homology arms of the AAVS1, FRT, SA, IRES, Zeocin-resistant gene, promoter (CAG, EF1α, UbC or CMV) GFP cDNA and right homology arm sequences into the XhoI and NotI restriction enzyme sites of the pBRMC1DTApAII plasmid. This was provided by Dr. Hitoshi Niwa (Kumamoto University).

Establishment of a cynomolgus monkey ES cell line

For establishment of a cynomolgus monkey embryonic stem cell line (Cyn ESC #3X), blastocysts were produced by intracytoplasmic sperm injection (ICSI) followed by *in vitro* culture for 8 days. ICSI and embryo culture were carried out as described by Seita *et al.* [26]. Inner cell masses (ICMs) were isolated by mechanical cutting using a needle. The ICMs were plated onto mouse embryonic fibroblast (MEF) feeder cells and cultured in embryonic stem (ES) cell medium as described below. After expanding ICMs, the primary colonies were dissected mechanically and transferred to feeder layers. After

several passages, the expanded cells were dissociated with TrypLE Select (Thermo Fisher Scientific 12563029) for further expansion. Karyotyping was carried out by Chromosome Science Labo Inc. (Sapporo, Japan), which showed that all 50 examined cells had the normal 42XX karyotype. For the analysis of *in vivo* differentiation ability, 5×10^5 ESCs were injected under the testicular capsules of severe combined immunodeficiency mice. Eight weeks after injection, teratomas were recovered and fixed with Bouin's fluid. Histology of prepared sections was performed using hematoxylin and eosin staining.

ES cell culture

Briefly, cynomolgus monkey ESCs were maintained on mitomycin C-treated MEF feeder cells in DMEM/F12 GlutaMax medium (Thermo Fisher Scientific, 10565-018) containing 10 ng/ml bFGF (Wako 6805384), 20% (v/v) KSR (Thermo Fisher Scientific, 10828028), NEAAs (Thermo Fisher Scientific, 11140-050), penicillin/streptomycin (Thermo Fisher Scientific 15140-122), 0.1 mM β-mercaptoethanol (Sigma-Aldrich M3148), and 10 μM XAV939 (Calbiochem 575545). For routine maintenance, cells were prepared for passaging every 5–7 days as single cell suspensions using TrypLE Select and seeded at a density of 1×10^4 cells/9 cm^2. The culture medium was supplemented with 10 μM ROCK inhibitor (Wako 253-00513) until 24 h after passaging. Two-hundred microliters of serum-free cell-freezing medium, BamBanker (Nippon Genetics CS-02-001), was used to cryopreserve aliquots of $1–5 \times 10^5$ cells.

To establish knock-in ESC lines, 2×10^4 ESCs were transfected simultaneously with circular forms of 1 μg targeting vector and 1 μg pX330 using Lipofectamine 2000 (Thermo Fisher Scientific). Transfected cells were seeded on MEFs. After 7–10 days

culture, GFP-positive primary colonies were picked out, dissociated with TrypLE Select and transferred onto fresh MEF layers in 48-well plates. When GFP-positive colonies did not appear, 96 colonies were picked up randomly, and PCR genotyping was carried out with primers as shown in Table S1 to obtain colonies in which the target gene was properly knocked in.

Embryoid body formation

After the passage period, 3,000 or 4,000 cells were suspended in EB medium (DMEM/F12 GlutaMax containing 20% (v/v) KSR, NEAA, Penicillin/Streptomycin, 0.1 mM β-Mercaptoethanol and 10 μM ROCK inhibitor), and aggregated in low-cell-binding V-bottom 96-well plates (Greiner Bio-one International 651970). The medium was not changed until the analysis at 8 days of induction. Images of the aggregates were taken using a Biorevo inverted microscope (Keyence). The aggregates were collected at the designated days for RNA, immunohistochemistry (IHC) and FACS analysis. For IHC, the aggregates were embedded in iPGell (GenoStaff), fixed with 4% paraformaldehyde (PFA), embedded in OCT compound, frozen in liquid nitrogen, cut into 10 μm-thick sections, and placed on glass slides that were treated with Blocking One (Nacalai Tesque) for 30 min at room temperature. Primary antibodies used are listed in Table S2. Labeled proteins were detected with appropriate secondary antibodies. Nuclei were counterstained with Hoechst 33342 (Thermo Fisher Scientific, H3570). Cells were then observed under a TCS SP8 confocal microscope (Leica).

For FACS analysis, after washing once with phosphate-buffered saline (PBS), the aggregates were dissociated by treatment with 0.25% trypsin-EDTA (Thermo Fisher Scientific) for 10–20 min at 37 °C and then dispersed by gentle pipetting. After washing with PBS containing 10% (v/v) fetal bovine serum (FBS, Sigma-Aldrich, 172012) and

0.1% bovine serum albumin (BSA, Sigma-Aldrich, A7906), the cells were suspended in FACS buffer (0.1% BSA in PBS) and passed through a cell strainer (BD Biosciences) to remove cell clumps. The numbers of cells per aggregate were counted and they were washed once with PBS and processed using a FACSCalibur instrument (BD Biosciences) for analysis and sorting.

GFP protein expression levels in CK8-positive endoderm cells or MSX1-positive mesoderm cells of EBs were analyzed as described [27]. Briefly, mean fluorescence intensities inside regions of interest were measured and subtracted from background signals, which were defined as the mean fluorescence intensities of randomly chosen cytoplasmic signals, and then normalized against the mean fluorescence intensities in the Hoechst channel using ImageJ ver. 1.51 image analysis software (National Institutes of Health, Bethesda, MD; https://imagej.nih.gov/ij/).

Neuronal differentiation

During the passage period, 3,000 or 4,000 cells were suspended in EB medium and aggregated in low-cell-binding V-bottom 96-well plates. On the second day after aggregation, the medium was changed from EB medium to neural differentiation medium. This had the same formulation as EB medium, with the addition of 1 μM retinoic acid (Wako 186-0114), 3 μM dorsomorphin (Wako 041-33763), 3 μM SB431542 (Wako 031-24291) and 3 μM BIO: 6-Bromoindirubin-3′-oxime (Wako 029-16241). After 4 days of aggregation, the medium was changed from neural differentiation medium to culture medium without dorsomorphin, SB431542 or BIO but with 1 μM purmorphamine (Wako 166-23991). After 14 days of differentiation, the aggregated cells were trypsinized into single cells and plated on Poly-D-Lysine (Sigma-Aldrich P1024)-coated glass slides and cultured for a further 5 days. After *in*

vitro neural differentiation, the cells were subjected to IHC with anti-SYNAPSIN I and anti-βIII-TUBULIN antibodies listed on Table S2.

Animals

All experimental procedures were approved by the Animal Care and Use Committee of Shiga University of Medical Science and methods were carried out in accordance with the approved guidelines (Approval number: 2016-12-5(H1)). Oocytes were collected from eight sexually mature female cynomolgus monkeys, aged 4–8 years and weighing 2.1–3.9 kg. Twenty sexually mature females aged 4 years old and weighing 2.0–3.8 kg, were used as recipients. Semen was collected from one sexually mature male monkey, aged 12 years and weighing 6.2 kg as described by Sankai et al. [28] with a minor modification. In brief, fresh spermatozoa were collected by direct electric stimulation (5–15 V, 1 pulse/sec) of the penis without anesthesia. For de-coagulation, ejaculated semen was placed in a disposable plastic tube and incubated at 37°C for 30 min. The semen was transferred into a new disposable plastic tube and diluted with 10 ml Biggers, Whitter, and Whittingham medium (BWW) containing 1 mM caffeine, 1 mM dbcAMP, and 0.3% BSA, and incubated at 37°C for 30 min. Then, the diluted semen was centrifuged at 380 × *g* for 5 min at room temperature to wash the semen. The supernatant was removed, and then 1 ml BWW containing 1 mM caffeine, 1 mM dbcAMP, and 0.3% BSA was added to the semen pellet. Spermatozoa that swam up from for the semen pellet were used for ICSI.

Temperature and humidity in the animal rooms were maintained at 25 ± 2 °C and 50 ± 5%, respectively. The light cycle was 12 h of artificial light from 08:00 to 20:00. In the morning, each animal was fed 20 g/kg of body weight of commercial pellet

monkey chow (CMK-1; CLEA Japan), supplemented with 20–50 g of sweet potato in the afternoon. Water was available *ad libitum*.

Lentiviral vector construction

pCSII-CAG-EGFP was constructed by introducing the CAG promoter from pCAGGS and GFP cDNA into pCSII-EF-MCS-IRES2-Venus plasmid. pCSII-EF1α-EGFP was constructed by introducing GFP cDNA into pCSII-EF-MCS-IRES2-Venus plasmid. pCAGGS was provided by Dr. Hitoshi Niwa (Kumamoto University).

Production of transgenic cynomolgus monkeys

Oocyte collection, lentiviral transduction, virus injection to embryos, ICSI, embryo transfer, pregnancy detection and observation of green fluorescence in Tg offspring were carried out as described by Seita *et al.* [26] The eight oocyte donors each received subcutaneous injection of human follicle-stimulating hormone (15 IU/kg, Asuka Pharmaceutical) via a micro-infusion pump (iPRECIO SMP-200, ALZET Osmotic Pumps) at 7 µl/h for 10 days. On day 10, the animals received an intramuscular injection of human chorionic gonadotropin (Puberogen, Nippon Zenyaku Kogyo), and oocytes were aspirated laparoscopically after 40 h. The recovered meiosis (M)-II-stage oocytes containing the first polar body were placed in m-TALP, a modified Tyrode's solution containing HEPES, and injected with lentivirus, ICSI was performed 3–4 h after virus injection. The fertilized oocytes were cultured in CMRL Medium-1066 (Thermo Fisher Scientific, 11530-037) containing 20% (v/v) FBS. When embryos developed to blastocysts, one embryo was transferred into each female recipient.

Southern blotting

Five microgram aliquots of genomic DNA were digested with EcoRI and the digested genomic DNA was separated on a 0.8% agarose gel and transferred to a Hybond-N+

nylon membrane (GE Healthcare Biosciences). Southern blot analysis was performed using the digoxigenin (DIG) system (Roche Diagnostics K.K.), according to the manufacturer's protocol. The GFP DNA probe was labeled by amplification in the presence of DIG using the PCR DIG probe synthesis kit.

Immunohistochemistry

Blastocysts were fixed in 4% PFA at room temperature for 30 min and treated with Blocking One (Nacalai Tesque) for 30 min at room temperature. Primary antibodies and dilutions used were listed on Table S2. GFP protein expression levels in the blastocysts were analyzed as described previously [27]. Briefly, mean fluorescence intensities inside regions of interest were measured and subtracted from background signals, which were defined as average of mean fluorescence intensities of random chosen cytoplasmic signals, and were then normalized against the mean fluorescence intensities in the Hoechst channel using ImageJ image analysis software as above

Tissues were fixed in 4% PFA at 4 °C overnight, embedded in OCT compound, frozen in liquid nitrogen, cut into 5- or 10-μm sections, placed onto glass slides and treated with Blocking One solution for 30 min at room temperature. Primary antibodies and dilutions were listed on Table S2. Labeled proteins were detected by appropriate secondary Alexa Fluor 594-labeled antibodies. Tissues were counterstained with Hoechst 33342 and the images were taken using the TCS SP8 confocal microscope.

RT-qPCR

Total RNA was extracted from cells or tissues using RNeasy Mini kits (Qiagen). For reverse transcription, ReverTra Ace (Toyobo TRT-101) and oligo (dT) 20 primers were used. THUNDERBIRD SYBR qPCR Mix (Toyobo QPS-101) was used for qPCR.

Transcript levels were determined in triplicate reactions and normalized against the corresponding levels of GAPDH. Primer sequences used are shown in Table S1.

Flow cytometry analysis

0.5 ml blood was collected from the femoral vein using a 27-gauge needle and centrifuged at 1,730 × *g* for 5 min to separate whole blood cells. Hemolysis was performed with Lysing buffer (BD Pharm. 555899) to collect mononuclear cells. These cells were washed with PBS and suspended in PBS + 2% (v/v) FBS. The pellet was incubated with mouse Alexa Fluor 647-conjugated anti-human CD20 (1:10, BioLegend 302318), mouse phycoerythrin-conjugated anti-human CD3 (1:10, BD 552127), and allophycocyanin-conjugated anti-mouse/human CD11b (1:100, BioLegend 101212) antibodies for 1 h on ice. Samples were washed with PBS and resuspended in 300 μl PBS containing 0.1 mg/ml propidium iodide. Flow cytometry analysis was then performed using a FACSCalibur instrument (BD Biosciences).

Statistical analysis

Statistical analyses of all data comparisons were carried out by multiple one way analysis of variance (ANOVA) using GraphPad Prism 8 software (https://www.graphpad.com/scientific-software/prism/). $P < 0.05$ was considered as statistically significant.

Results

Evaluation of various ubiquitous promoters in undifferentiated ESCs

Although previous studies reported the generation of transgenic NHPs that ubiquitously overexpress a transgene such as that encoding GFP by utilizing various promoters [20–22,29], it is not clear which promoter is most suitable for persistent and ubiquitous overexpression in undifferentiated and differentiated tissues of NHPs. Therefore, we

aimed to clarify suitable promoters for ubiquitous and stable Tg overexpression in the tissues of cynomolgus monkeys. In particular, we paid attention to the level and extent of expressions driven by a promoter and its persistent expression (i.e., resistance to the silencing of Tg expression).

Although we have generated GFP-expressing Tg cynomolgus monkeys and showed that the CAG promoter drives transgene expression in most cells of various tissues [26], there are inherent difficulties in the evaluation of promoter activities using such models; thus, the copy number of the lentiviral vector and its integration site into the genome are not easily controlled, and animal experiments must be restricted because of the high ethical standards required by NHP studies. Therefore, we aimed to develop an *in vitro* assay system to evaluate promoter activities in undifferentiated and differentiated cynomolgus monkey ESCs. To compare promoter activities under the same experimental condition, we tried to introduce a promoter-GFP knock-in donor into the first intron of the AAVS1 locus, which is often used as a safe harbor [30–32], in cynomolgus monkey ESCs (Fig.1A), thereby enabling us to measure promoter activity under identical chromosome positions and copy numbers. First, we established a novel cynomolgus monkey ESC line (Fig. 1B) and checked the karyotyping by G-banding and pluripotency by teratoma formation ability (Fig. 1C, D). PCR analysis showed that AAVS1 locus were correctly targeted with respective promoter-GFP knock-in vector, and the heterozygous knock-in ESC lines were established (Fig.1E). We examined GFP protein expression and found the strongest GFP fluorescence in EF1α-GFP ESCs under undifferentiated condition (Fig.1F), whereas the CAG and UbC promoters were second and third strongest, respectively (Fig.1F). Unexpectedly, the CMV promoter showed no detectable activity (Fig. 1F). Fluorescence-activated cell sorting (FACS) and RT–qPCR

analysis also showed the strongest GFP expression was driven by the EF1α promoter (Fig. 1G–I), although there were GFP-negative cells (24.3%) in EF1α-GFP ESCs (Fig. 1G). Because the CMV promoter was unexpectedly inactive, we introduced the promoter-GFP knock-in donor vectors used for targeting experiments into 293FT cells transiently and found that the CMV promoter was highly active in this situation (Fig. S1).

Taken together, the EF1α promoter proved most suitable for a high level of overexpression in undifferentiated pluripotent stem cells; the CAG and UbC promoters were suitable as second and third choices, respectively.

Evaluation of promoters in embryoid bodies (EBs) harboring three germ layer-derived tissues

To evaluate promoter activities in differentiated cells, EBs harboring three germ layer-derived tissues were prepared in floating culture in the absence of basic fibroblast growth factor (bFGF) (Fig. 2A) as described [33]. ESCs formed ball-shaped structures in the suspension culture for 8 days. Semiquantitative PCR analysis indicated that the pluripotency markers, *OCT4* and *NANOG*, decreased rapidly, and the endoderm markers, *CK8* and *CK18*, the mesoderm markers, *T* (*BRACHYURY)* and *MSX1*, and ectoderm markers, *PAX6* and *MAP2*, increased, indicating that the EBs had differentiated successfully into all three germ layers (Fig. 2B). AAVS1$^{+/CAG\text{-}GFP}$ (AAVS1-CAG-GFP) EBs at day 8 showed the strongest fluorescence, and AAVS1$^{+/EF1\alpha\text{-}GFP}$ (AAVS1-EF1α-GFP) EBs and AAVS1$^{+/UbC\text{-}GFP}$ (AAVS1-UbC-GFP) EBs showed moderate fluorescence. No fluorescence was observed in AAVS1$^{+/CMV\text{-}GFP}$ (AAVS1-CMV-GFP) EBs (Fig. 2C). FACS analysis showed a high percentage of GFP-positive cells in AAVS1-CAG-GFP (94.5%) and AAVS1-UbC-GFP (98.9%) EBs

at day 8, compared with those of AAVS1-EF1α-GFP EBs (72.8%) and AAVS1-CMV-GFP EBs (0.57%) (Fig. 2D). The fluorescence intensity of AAVS1-CAG-GFP EBs was 1.3 times stronger than that of AAVS1-EF1α-GFP and 2.3 times stronger than that of AAVS1-UbC-GFP EBs at day 8 of culture (Fig. 2E). RT–qPCR analysis showed that GFP mRNA expression in AAVS1-CAG-GFP EBs increased significantly as the tissues differentiated, while GFP mRNA expression in AAVS1-EF1α-GFP EBs decreased remarkably (Fig. 2F). However, the GFP mRNA expression level in AAVS1-UbC-GFP EBs showed no significant change (Fig. 2F). We also examined GFP expression levels in *CK8*-positive endoderm cells and *MSX1*-positive mesoderm cells in the EBs at day 8 (Figs. 2G, H, S2). We found that the CAG promoter had strong and uniform activity in both *CK8*- and *MSX*-positive cells (Figs. 2G, H, S2), while the EF1α promoter showed bimodal expressions in these cells; strong and weak/negative, respectively. The UbC promoter had moderate and uniform activity in both *CK8*- and *MSX1*-positive cells (Figs. 2G, H, S2), while the CMV promoter activity was very weak in these cells.

Promoter activities in ESC-derived neurons

To evaluate promoter activities in ES-derived neurons, ESCs were differentiated into neurons according to the dual inhibition method [34] (Fig. 3A). RT–qPCR and immunohistochemistry (IHC) showed that βIII TUBULIN and SYNAPSIN I expressions were induced during the course of EB differentiation (Fig. 3B), indicating successful differentiation of the ESCs into neurons. AAVS1-CAG-GFP ESCs showed strongest GFP fluorescence (Fig. 3C), while AAVS1-EF1α-GFP and AAVS1-UbC-GFP ESCs showed moderate signals, and AAVS1-CMV-GFP ESCs showed no detectable

fluorescence (Fig. 3C). Consistent with these observations, the mean intensities of GFP-positive cells in FACS analysis of the EBs at day 14 and the results of RT–qPCR analysis showed that the CAG promoter showed the strongest signal (Fig. 3D, E). FACS analysis showed that most if not all of the βIII TUBULIN-positive neurons in AAVS1-CAG-GFP EBs were strongly GFP-positive (98.1%)(Fig. 3F), while AAVS1-EF1α-GFP EBs were GFP-positive (65.8%) with a substantial proportion of negative cells (33.5%), AAVS1-UbC-GFP EBs were GFP-positive (93.1%), but the intensity was weaker than that of AAVS1-CAG-GFP EBs, AAVS1-CMV-GFP EBs were almost entirely GFP-negative (99.6%) (Fig. 3F). At 19 days, most of the SYNAPSIN I-positive neurons from AAVS1-CAG-GFP ESCs co-expressed GFP strongly (Fig. S3), while SYNAPSIN I-positive neurons from AAVS1-EF1α-GFP ESCs showed bimodal GFP expression with GFP-negative cells (Fig. S3; the cells indicated by arrows). Thus, the CAG promoter proved the most suitable for persistent and high level transgenic expression in differentiated cells.

Production of GFP Tg cynomolgus monkeys with CAG and EF1α promoters

Because the CAG and EF1α promoters drove transgenic expression most strongly in differentiated and undifferentiated cells, respectively, we constructed lentiviral vectors carrying the gene encoding GFP under the control of these promoters (Fig. 4A) and injected them into oocytes. Eight days after fertilization, the blastocysts infected with EF1α-GFP lentivirus showed stronger GFP fluorescence than did those treated with the CAG-GFP lentivirus (Fig. 4B), consistent with the notion that the EF1α promoter drives stronger Tg expression in undifferentiated ESCs than does the CAG promoter (Fig. 1F–I). After transfer to recipient foster mothers, we obtained three CAG-GFP Tg and three

EF1α-GFP Tg offspring (Fig. 4C, Table S3). We estimated the copy numbers of transgenes by Southern blot analysis as follows (Fig. S4): CAG-GFP Tg monkeys (CE1894M, 1 copy; CE1993F, 5 copies; CE1984F, 5 copies) and EF1α-GFP Tg monkeys (CE1881M 4 copies; CE1886M 11 copies; and CE1887F, 8 copies). GFP fluorescence was observed on the facial skin of the CAG-GFP CE1894M, CE1993F and CE1984F Tg offspring, while GFP fluorescence was not detected on the facial skin of the EF1α-GFP Tg monkey (CE1881M), and GFP fluorescence was observed on the facial skin of the CE1886M and CE1887F. It is notable that GFP fluorescence was detected in a CAG-GFP Tg monkey carrying just one copy of the gene for GFP (CE1894M), whereas it was not detected in the EF1α-GFP Tg monkey carrying four copies of the transgenes (CE1881M) (Figs. 4C, S4). Consistent with our finding that the CAG promoter drove transgenic expression more strongly than the EF1α promoter in ESC-derived differentiated cells (Figs. 2, 3, S3), activity of the CAG promoter was overall stronger in the skin than that of the EF1α promoter (Fig. 4).

GFP expression in the skin tissues of Tg animals

Although GFP fluorescence intensity of facial skin in the CAG-GFP Tg monkey was overall stronger than that of the EF1α-GFP Tg monkey, we investigated the expression of GFP at a cellular level in back skin biopsies (Fig. 5). IHC showed that GFP expression could be detected in the hair roots of three CAG-GFP Tg offspring (CE1894M, CE1993F, and CE1984F), and that the intensity increased with the transgene copy number (Figs. 5A, S4). However, the EF1α-GFP Tg, monkey (CE1881M) did not show detectable GFP signals in the hair follicle, even though it carried four copies of the transgene (Figs 5A, S4). The CE1886M and CE1887F showed

small numbers of GFP-positive cells in the hair follicles (Figs. 5A, S4). Consistent with the GFP protein expression in skin tissues, RT–qPCR analysis showed that GFP mRNA expression was abundant in the skin of CAG-GFP Tg monkeys but at a low level in the EF1α-GFP Tg monkeys (Fig. 5B). Overall, these data showed that GFP expression levels in the skin tissues of CAG-GFP Tg monkeys were considerably stronger than in EF1α-GFP Tg monkeys, although the exact copy numbers and chromosomal integration sites of the transgenes were not controlled precisely.

GFP expression in the blood cells of Tg animals

Since we investigated promoter activities in the skin that is derived from ectoderm, we investigated peripheral blood cells originating from mesoderm. FACS analysis clearly showed that the percentage of GFP-positive cells in T-cells (CD3+), B-cells (CD20+), granulocytes (CD11+/SSC high) and monocytes (CD11+/SSC low) were higher than those of EF1α-GFP Tg cells in CAG-GFP Tg monkeys (Fig. S5). The percentage and fluorescence intensity of GFP-positive peripheral blood cells tended to increase as the GFP transgene copy number increased, in both CAG-GFP Tg and EF1α-GFP Tg monkeys (Fig. S5). Taken together, the CAG promoter proved more suitable to drive transgene expression in hematopoietic cells at high levels than the EF1α promoter.

In summary, our results indicate that the CAG promoter is most suitable for stable, ubiquitous and high-level expression of transgenes in ESC-derived differentiated cells and tissues of cynomolgus monkeys (Fig. 6). The EF1α promoter is suitable for high expression of transgenes in undifferentiated ESCs and blastocysts, but it undergoes silencing (Fig. 6). The UbC promoter is suitable for stable, ubiquitous, and moderate expression of transgenes in ESC-derived differentiated cells (Fig. 6).

Discussion

Because NHPs share similar developmental paths with humans in their anatomy, physiology and genetics [35], there has been growing interest in human disease models using genetically modified NHPs [21,36,37]. However, it is not yet clear which promoter is most suitable to create Tg animals with ubiquitous and persistent overexpression of a transgene in their tissues. Given the fundamental ethical dilemma concerning the use of NHPs in biomedical research [35], we tried to substitute animal experiments with *in vitro* ESC differentiation. The AAVS1 locus of cynomolgus monkey ESCs was targeted with four well-known promoters: CAG, EF1α, CMV and UbC, which are frequently used in other animal species [38–41]. Our study clearly demonstrated that the EF1α promoter drove the strongest Tg expression in undifferentiated cynomolgus monkey ESCs, and that the CAG and UbC promoters were second and third in efficacy, respectively. This finding was consistent with a previous report by Xia *et al*., where the descending order of promoter activity was EF1α, CAG, and CMV [42] in undifferentiated human ESCs. Our study also clearly demonstrated that CAG promoter drove Tg expression ubiquitously at the highest level in EBs and ES-derived neurons without detectable epigenetic silencing, consistent with the report by Jakobsson *et al*., that the CAG promoter among three tested (CMV, EF1α and CAG) achieved a high level of GFP expression in the striatum and white matter of the brain [43]. The EF1α promoter drove transgene expression at the highest level in most of undifferentiated ESCs, but lost its activity in some cells. Potentially, an epigenetic silencing mechanism, which was proposed previously, occurred in our ES cells and tissues of Tg animals [26], although the exact mechanism is unclear. As ESCs differentiated, the Tg expression was reduced

gradually, consistent with a report that the EF1α promoter suffers from gradual silencing in mouse and human ESCs [44,45]. Our finding that GFP expression was not detectable in EF1α-GFP Tg monkeys with four copies of the transgene might have arisen in part from silencing of the corresponding genomic regions is inherently difficult because the blastocysts did not use the safe harbor approach.

Mosaicism is an issue to be considered when Tg animals are generated by virus-mediated approaches [46,47], because viral DNA may be integrated into the genome after the first cleavage, even though the virus infects fertilized eggs. Such mosaicism may in part explain our results, i.e. the differences in the band intensities of our Southern blot and the GFP-negative cells in Tg monkey tissues.

Transgenic animals ubiquitously overexpressing a transgene (e.g., that encoding GFP) have been widely used in biomedical research including bone marrow transplantation. In general, a line of Tg animals carrying a single integration of a transgene in a chromosome is used, because the line can be expanded by mating, and the level and pattern of GFP expression in various tissues is identical—in principle—among the individuals suitable for experimental reproducibility. In this regard, the CAG-GFP Tg monkey should be a promising founder animal for future generations, because even a single copy of the CAG-GFP transgene could label most if not all cells in various tissues, while other Tg monkeys carry multiple copies of transgenes and each transgene needs to be segregated. It is of note that a single copy of the EF1α-GFP transgene seems to be insufficient to drive GFP expression in skin and hematopoietic tissues.

Taken together, our study clearly demonstrates that the CAG promoter is the best option to achieve ubiquitous and stable expression of transgene in the tissues of cynomolgus monkeys used for human disease modeling. We also suggest that *in vitro*

ESC differentiation systems can substitute significantly for animal experiments using NHPs.

Funding

This study was supported in part by JSPS KAKENHI Grant Numbers JP17K14977 and JP17K08135 to Y.S. and T.T., respectively, PRESTO, Japan Science and Technology Agency to M.E., and by an in-house grant from Shiga University of Medical Science to M.E., Y.S., and T.T.

Acknowledgements

We thank the Research Center for Animal Life Science research support team for animal care at Shiga University of Medical Science. We thank James Cummins, PhD, from EDANZ Group (www.edanzediting.com/ac) for editing a draft of this manuscript.

Author contributions

Y. S. performed the biochemical and cell biological experiments and co-wrote the paper. Y. S. and M. E. designed the experiments. T. T., Y. S. and H. M designed the GFP expression vectors and prepared lentivirus particles. C. I., H. T., S. N., and Y. S. assisted in embryological technique development. T. T. established ESCs. T. T., K. K., M. N. and A. K performed teratoma assays. T. T. and S. H assisted with ESC differentiation. T. A. and Y. S. contributed to immunohistochemistry. Y. S. performed FACS analysis. H. T. and Y. S. performed Southern blot analysis. T. A., T. T., and Y. S. analyzed the data. M. E. conducted the project, contributed to the development of hypotheses and co-wrote the paper.

Competing financial interests

The authors declare no competing financial interests.

References

[1] Rosenthal N, Brown S. The mouse ascending : perspectives for human-disease models. Nat Cell Biol 2007; 9:993–999.

[2] Gordon JW, Scangos GA, Plotkin DJ, Barbosa JA, Ruddle FH. Genetic transformation of mouse embryos by microinjection of purified DNA. Proc Natl Acad Sci U S A 1980; 77:7380–7384.

[3] Gossler A, Doetschman T, Korn R, Serflingt E, Kemler R. Transgenesis by means of blastocyst-derived embryonic stem cell lines. Proc Natl Acad Sci U S A 1986; 83:9065–9069.

[4] Robertson E, Bradley A, Kuehn M, Evans M. Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. Nature 1986; 323:445–448.

[5] Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R. One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. Cell 2013; 153:910–918.

[6] Sung YH, Baek I-J, Kim DH, Jeon J, Lee J, Lee K, Jeong D, Kim J-S, Lee H-W. Knockout mice created by TALEN-mediated gene targeting. Nat Biotechnol 2013; 31:23–24.

[7] Gautier CA, Kitada T, Shen J. Loss of PINK1 causes mitochondrial functional defects and increased sensitivity to oxidative stress. Proc Natl Acad Sci U S A 2008; 105:11364–11369.

[8] Andres-Mateos E, Perier C, Zhang L, Blanchard-Fillion B, Greco TM, Thomas B, Ko HS, Sasaki M, Ischiropoulos H, Przedborski S, Dawson TM, Dawson VL. DJ-1

gene deletion reveals that DJ-1 is an atypical peroxiredoxin-like peroxidase. Proc Natl Acad Sci U S A 2007; 104:14807–14812.

[9] Hinkle KM, Yue M, Behrouz B, Dächsel JC, Lincoln SJ, Bowles EE, Beevers JE, Dugger B, Winner B, Prots I, Kent CB, Nishioka K, et al. LRRK2 knockout mice have an intact dopaminergic system but display alterations in exploratory and motor co-ordination behaviors. Mol Neurodegener 2012; 7:25.

[10] Itier J-M, Ibanez P, Mena MA, Abbas N, Cohen-Salmon C, Bohme GA, Laville M, Pratt J, Corti O, Pradier L, Ret G, Joubert C, et al. Parkin gene inactivation alters behaviour and dopamine neurotransmission in the mouse. Hum Mol Genet 2003; 12:2277–2291.

[11] Kitada T, Pisani A, Karouani M, Haburcak M, Martella G, Tscherter A, Platania P, Wu B, Pothos EN, Shen J. Impaired dopamine release and synaptic plasticity in the striatum of parkin-/- mice. J Neurochem 2009; 110:613–621.

[12] Hennis MR, Marvin MA, Taylor CM, Goldberg MS. Surprising behavioral and neurochemical enhancements in mice with combined mutations linked to Parkinson's disease. Neurobiol Dis 2014; 62:113–123.

[13] Breschi A, Gingeras TR, Guigó R. Comparative transcriptomics in human and mouse. Nat Rev Genet 2017; 18:425–440.

[14] Saitou M, Miyauchi H. Gametogenesis from Pluripotent Stem Cells. Cell Stem Cell 2016; 18:721–735.

[15] Rossant J, Tam PPL. New Insights into Early Human Development: Lessons for Stem Cell Derivation and Differentiation. Cell Stem Cell 2017; 20:18–28.

[16] Lane MA. Nonhuman primate models in biogerontology. Exp Gerontol 2000; 35:533–541.

[17] King MC, Wilson AC. Evolution at two levels in humans and chimpanzees. Science 1975; 188:107–116.

[18] McConkey EH, Varki A. A primate genome project deserves high priority. Science 2000; 289:1295–1296.

[19] Hayakawa T, Suzuki-Hashido N, Matsui A, Go Y. Frequent Expansions of the Bitter Taste Receptor Gene Repertoire during Evolution of Mammals in the Euarchontoglires Clade. Mol Biol Evol 2014; 31:2018–2031.

[20] Chan AW, Chong KY, Martinovich C, Simerly C, Schatten G. Transgenic monkeys produced by retroviral gene transfer into mature oocytes. Science 2001; 291:309–312.

[21] Yang S-H, Cheng P-H, Banta H, Piotrowska-Nitsche K, Yang J-J, Cheng ECH, Snyder B, Larkin K, Liu J, Orkin J, Fang Z-H, Smith Y, et al. Towards a transgenic model of Huntington's disease in a non-human primate. Nature 2008; 453:921–924.

[22] Sasaki E, Suemizu H, Shimada A, Hanazawa K, Oiwa R, Kamioka M, Tomioka I, Sotomaru Y, Hirakawa R, Eto T, Shiozawa S, Maeda T, et al. Generation of transgenic non-human primates with germline transmission. Nature 2009; 459:523–527.

[23] Wolfgang MJ, Eisele SG, Browne MA, Schotzko ML, Garthwaite MA, Durning M, Ramezani A, Hawley RG, Thomson JA, Golos TG. Rhesus monkey placental transgene expression after lentiviral gene transfer into preimplantation embryos. Proc Natl Acad Sci U S A 2001; 98:10728–10732.

[24] Tomioka I, Nogami N, Nakatani T, Owari K, Fujita N, Motohashi H, Takayama O, Takae K, Nagai Y, Seki K. Generation of transgenic marmosets using a

tetracyclin-inducible transgene expression system as a neurodegenerative disease model. Biol Reprod 2017; 97:772–780.

[25] Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. Multiplex Genome Engineering Using CRISPR/Cas Systems. Science 2013; 339:819–823.

[26] Seita Y, Tsukiyama T, Iwatani C, Tsuchiya H, Matsushita J, Azami T, Okahara J, Nakamura S, Hayashi Y, Hitoshi S, Itoh Y, Imamura T, et al. Generation of transgenic cynomolgus monkeys that express green fluorescent protein throughout the whole body. Sci Rep 2016; 6:24868.

[27] Azami T, Waku T, Matsumoto K, Jeon H, Muratani M, Kawashima A, Yanagisawa J, Manabe I, Nagai R, Kunath T, Nakamura T, Kurimoto K, et al. Klf5 maintains the balance of primitive endoderm versus epiblast specification during mouse embryonic development by suppression of Fgf4. Development 2017; 144:3706–3718.

[28] Sankai T, Terao K, Yanagimachi R, Cho F, Yoshikawa Y. Cryopreservation of spermatozoa from cynomolgus monkeys (Macaca fascicularis). J Reprod Fertil 1994; 101:273–278.

[29] Niu Y, Yu Y, Bernat A, Yang S, He X, Guo X, Chen D, Chen Y, Ji S, Si W, Lv Y, Tan T, et al. Transgenic rhesus monkeys produced by gene transfer into early-cleavage-stage embryos using a simian immunodeficiency virus-based vector. Proc Natl Acad Sci U S A 2010; 107:17663–17667.

[30] Smith JR, Maguire S, Davis LA, Alexander M, Yang F, Chandran S, Ffrench-Constant C, Pedersen RA. Robust, persistent transgene expression in

human embryonic stem cells is achieved with AAVS1-targeted integration. Stem Cells 2008; 26:496–504.

[31] Luo Y, Liu C, Cerbini T, San H, Lin Y, Chen G, Rao MS, Zou J. Stable Enhanced Green Fluorescent Protein Expression After Differentiation and Transplantation of Reporter Human Induced Pluripotent Stem Cells Generated by AAVS1 Transcription Activator-Like Effector Nucleases. Stem Cells Transl Med 2014; 3:821–835.

[32] Lombardo A, Cesana D, Genovese P, Di Stefano B, Provasi E, Colombo DF, Neri M, Magnani Z, Cantore A, Lo Riso P, Damo M, Pello OM, et al. Site-specific integration and tailoring of cassette design for sustainable gene transfer. Nat Methods 2011; 8:861.

[33] Itskovitz-Eldor J, Schuldiner M, Karsenti D, Eden A, Yanuka O, Amit M, Soreq H, Benvenisty N. Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers. Mol Med 2000; 6:88–95.

[34] Morizane A, Doi D, Kikuchi T, Nishimura K, Takahashi J. Small-molecule inhibitors of bone morphogenic protein and activin/nodal signals promote highly efficient neural induction from human pluripotent stem cells. J Neurosci Res 2011; 89:117–126.

[35] Phillips KA, Bales KL, Capitanio JP, Conley A, Czoty PW, Hopkins WD, Hu S-L, Miller LA, Nader MA, Nathanielsz PW, Rogers J, Shively CA, et al. Why Primate Models Matter. Am J Primatol 2014; 76:801-827

[36] Niu Y, Guo X, Chen Y, Wang C, Gao J, Yang W, Kang Y, Si W, Wang H, Yang S, Li S, Ji W, et al. Early Parkinson ’ s disease symptoms in α -synuclein transgenic monkeys 2015; 24:2308–2317.

[37] Kang Y, Zheng B, Shen B, Chen Y, Wang L, Wang J, Niu Y, Cui Y, Zhou J, Wang H, Guo X, Hu B, et al. CRISPR / Cas9-mediated Dax1 knockout in the monkey recapitulates human AHC-HH. Hum Mol Genet 2015: 24:7255–7264.

[38] Okabe M, Ikawa M, Kominami K, Nakanishi T, Nishimune Y. 'Green mice' as a source of ubiquitous green cells. FEBS Lett 1997; 407:313–319.

[39] Binkley N, Kimmel D, Bruner J, Haffa A, Davidowitz B, Meng C, Schaffer V, Green J. Zoledronate prevents the development of absolute osteopenia following ovariectomy in adult rhesus monkeys. J Bone Miner Res 1998; 13:1775–1782.

[40] Masson D, Deckert V, Gautier T, Klein A, Desrumaux C, Viglietta C, Pais de Barros J-P, Le Guern N, Grober J, Labbe J, Menetrier F, Ripoll P-J, et al. Worsening of Diet-Induced Atherosclerosis in a New Model of Transgenic Rabbit Expressing the Human Plasma Phospholipid Transfer Protein. Arterioscler Thromb Vasc Biol 2011; 31:766–774.

[41] Dholakia U, Bandyopadhyay S, Hod EA, Prestia KA. Determination of RBC Survival in C57BL/6 and C57BL/6-Tg(UBC-GFP) Mice. Comp Med 2015; 65:196–201.

[42] Xia X, Zhang Y, Zieth CR, Zhang S-C. Transgenes delivered by lentiviral vector are suppressed in human embryonic stem cells in a promoter-dependent manner. Stem Cells Dev 2007; 16:167–176.

[43] Jakobsson J, Ericson C, Jansson M, Björk E, Lundberg C. Targeted transgene expression in rat brain using lentiviral vectors. J Neurosci Res 2003; 73:876–885.

[44] Norrman K, Fischer Y, Bonnamy B, Wolfhagen Sand F, Ravassard P, Semb H. Quantitative comparison of constitutive promoters in human ES cells. PLoS One 2010; 5:e12413.

[45] Hong S, Hwang D-Y, Yoon S, Isacson O, Ramezani A, Hawley RG, Kim K-S. Functional analysis of various promoters in lentiviral vectors at different stages of in vitro differentiation of mouse embryonic stem cells. Mol Ther 2007; 15:1630–1639.

[46] Pfeifer A. Lentiviral Transgenesis - A Versatile Tool for Basic Research and Gene Therapy. Curr Gene Ther 2006; 6:535–542.

[47] Wongsrikeao P, Saenz D, Rinkoski T, Otoi T, Poeschla E. Antiviral restriction factor transgenesis in the domestic cat. Nat Methods 2011; 8:853–859.

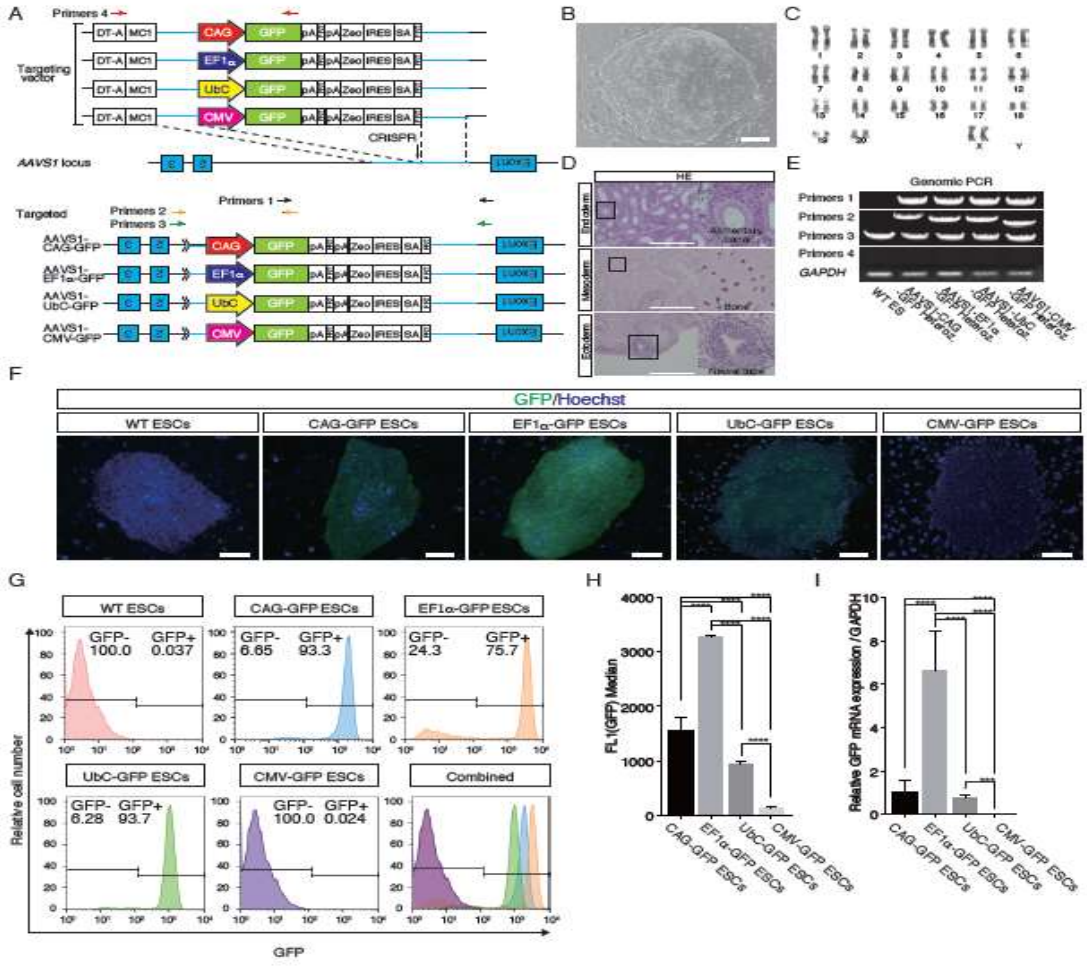

Figure 1. Generation of AAVS1-CAG-GFP/AAVS1-EF1α-GFP/AAVS1-UbC-GFP/AAVS1-CMV-GFP ESCs

A. Schematic overview depicting the targeting strategy for the cynomolgus monkey *AAVS1* locus. Arrow, genomic site cut by the respective CRISPR pair. Shown above is a schematic of the donor plasmid design. Donor plasmids were created corresponding to the cleavage location of the CRISPR pair and transferred roughly 1,500-bp regions of homology to the *AAVS1* genomic sequence. Red, black, orange, and green arrows show primer pairs for genomic PCR. Each primer pair amplified the corresponding genomic regions: primer pair 1, 3′ half of the knock-in construct, 3′

arm, and a part of the genomic region outside of the 3′ arm; primer pair 2, 5′ half of the knock-in construct, 5′ arm, and a part of the genomic region outside of the 5′ arm; primer pair 3, the entire knock-in construct, 5′ and 3′ arms, and a part of the genomic region outside of the 5′ and 3′ arms; primer pair 4, a part of the DT-A and MC1 promoter and 3′ half of the knock-in construct. SA, splice acceptor sequence; *Zeo*, zeocin resistance gene; polyA, polyadenylation sequence; FRT, flippase recognition target sequence; IRES, internal ribosome entry site; DT-A, Diphtheria toxin A sequence. **B.** Establishment of cynomolgus monkey ESC line. **C.** Karyotyping of ESCs. **D.** Teratoma derived from ESCs. HE, Hematoxylin and eosin (HE) staining of a teratoma derived from ESCs (Cyn ESC #3X). Scale bars = 300 μm. **E.** Genomic PCR of AAVS1-CAG-GFP, AAVS1-EF1α-GFP, AAVS1-UbC-GFP and AAVS1-CMV-GFP ESCs. WT, wild-type. **F.** Fluorescent imaging of GFP in heterozygous AAVS1-CAG-GFP, AAVS1-EF1α-GFP, AAVS1-UbC-GFP and AAVS1-CMV-GFP ESCs. Scale bars = 100 μm. **G.** FACS analysis of heterozygous AAVS1-CAG-GFP, AAVS1-EF1α-GFP, AAVS1-UbC-GFP and AAVS1-CMV-GFP ESCs. **H.** Median GFP fluorescence intensity of heterozygous AAVS1-CAG-GFP, AAVS1-EF1α-GFP, AAVS1-UbC-GFP and AAVS1-CMV-GFP ESCs. Data are shown as the mean of median intensities. **I.** RT–qPCR analysis of heterozygous AAVS1-CAG-GFP, AAVS1-EF1α-GFP, AAVS1-UbC-GFP and AAVS1-CMV-GFP ESCs. Data are shown as the mean ± standard deviation (SD). Asterisks in **H** and **I** indicate the statistical significance of differences: ****$P < 0.0001$; ***$P < 0.001$.

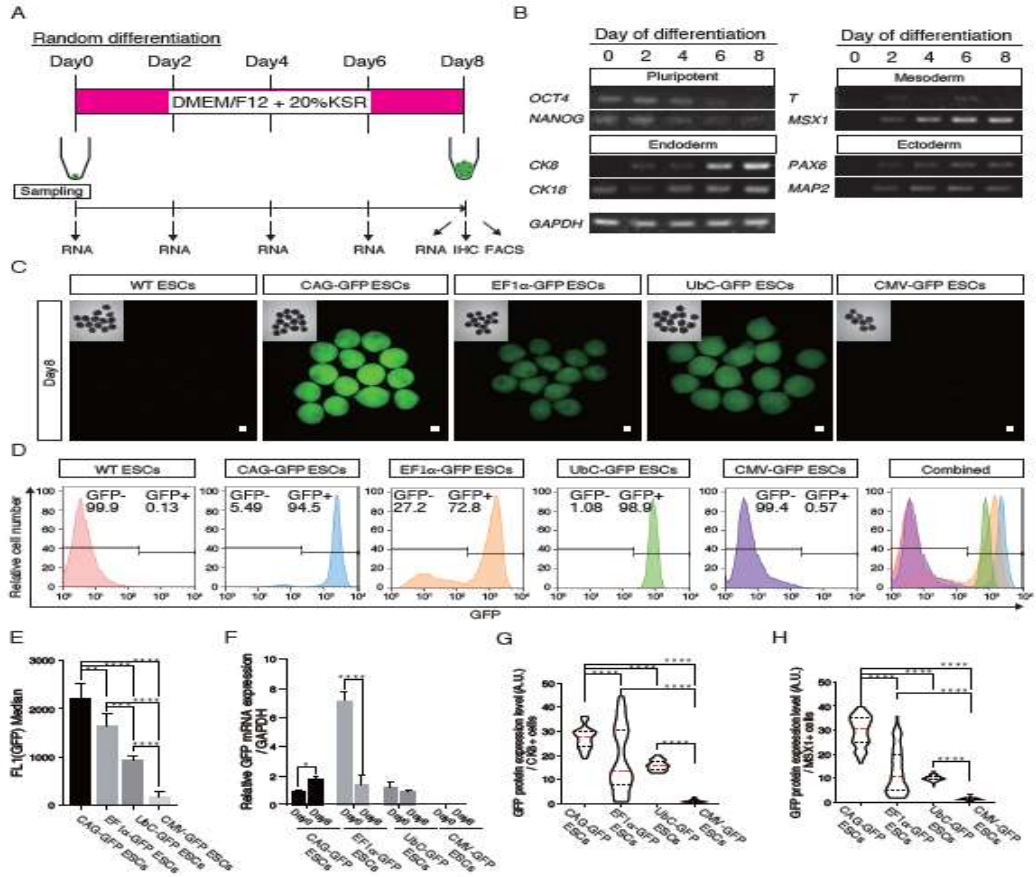

Figure 2. Promoter activities in embryoid bodies with all three germ layers

A. Schematic overview for the random differentiation of ESCs. Shown above is the medium for random differentiation of ESCs. Shown below is the timing of cell sampling. **B.** Semiquantitative PCR analyses of various differentiation markers for three germ layers. **C.** Fluorescent images of day 8 EBs from heterozygous AAVS1-CAG-GFP, AAVS1-EF1α-GFP, AAVS1-UbC-GFP and AAVS1-CMV-GFP ESCs. Insets in each panel show brightfield images. Scale bars = 100 μm. **D.** FACS analysis using day 8 EBs from heterozygous AAVS1-CAG-GFP, AAVS1-EF1α-GFP, AAVS1-UbC-GFP and AAVS1-CMV-GFP ESCs. **E.** GFP fluorescence intensities in day 8 EBs from heterozygous AAVS1-CAG-GFP, AAVS1-EF1α-GFP, AAVS1-UbC-GFP and AAVS1-CMV-GFP ESCs. Data are shown as means of the

median. **F.** GFP expression levels in day 0 and day 8 EBs from heterozygous AAVS1-CAG-GFP, AAVS1-EF1α-GFP, AAVS1-UbC-GFP and AAVS1-CMV-GFP ESCs evaluated by RT–qPCR. Data are shown as the mean ± SD. **G.** Violin plot of GFP protein expression level of CK8-positive cells in the EBs at day 8. The red dashed line indicates the median value and the black dashed lines indicate quartiles. **H.** Violin plot of GFP protein expression level of MSX1-positive cells in the EBs at day 8. The red dashed line indicates the median value and the black dashed lines indicate quartiles. Asterisks in **E, F, G,** and **H** indicate statistical significance: ****$P < 0.0001$; ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$

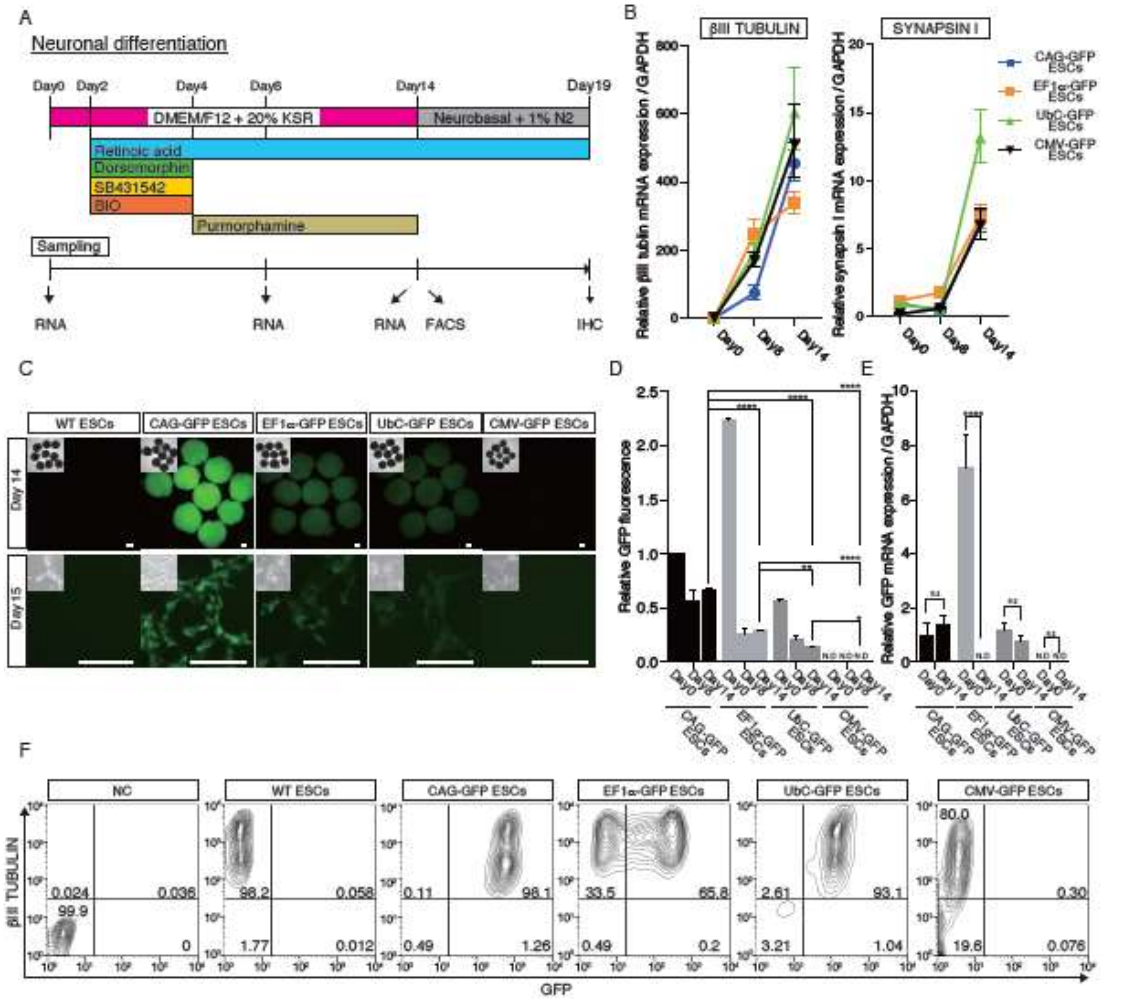

Figure 3. Promoter activities in the ESC-derived neurons

A. Schematic overview for the neuronal differentiation of ESCs. Shown above is the medium for neuronal differentiation of ESCs. Shown below is the timing of cell sampling. **B.** Expression of neuronal maker genes during differentiation measured by RT–qPCR. **C.** Fluorescent images of day 14 EBs (upper) and dispersed cells cultured for 1 day after trypsinization of the EBs (lower) formed by heterozygous AAVS1-CAG-GFP, AAVS1-EF1α-GFP, AAVS1-UbC-GFP and AAVS1-CMV-GFP ESCs. Insets in each panel shows brightfield images. Scale bars = 100 μm. **D.** Relative GFP fluorescence intensities in day 0, day 8 and day 14 EBs from heterozygous AAVS1-CAG-GFP,

AAVS1-EF1α-GFP, AAVS1-UbC-GFP and AAVS1-CMV-GFP ESCs. Fluorescence was measured by FACS analysis. Data are shown as mean ± SD. **E.** GFP expressions in day 0 and day 14 EBs from heterozygous AAVS1-CAG-GFP, AAVS1-EF1α-GFP, AAVS1-UbC-GFP and AAVS1-CMV-GFP ESCs evaluated by RT-qPCR. Data are represented as mean ± SD. N.D., Not detected. Asterisks in **D.** and **E.** indicate statistical significance: ****$P < 0.0001$; **$P < 0.01$; *$P < 0.05$ **F.** FACS analyses of βIII TUBULIN/GFP expression of the day 14 EBs from heterozygous AAVS1-CAG-GFP, AAVS1-EF1α-GFP, AAVS1-UbC-GFP and AAVS1-CMV-GFP ESCs. WT, wild-type.

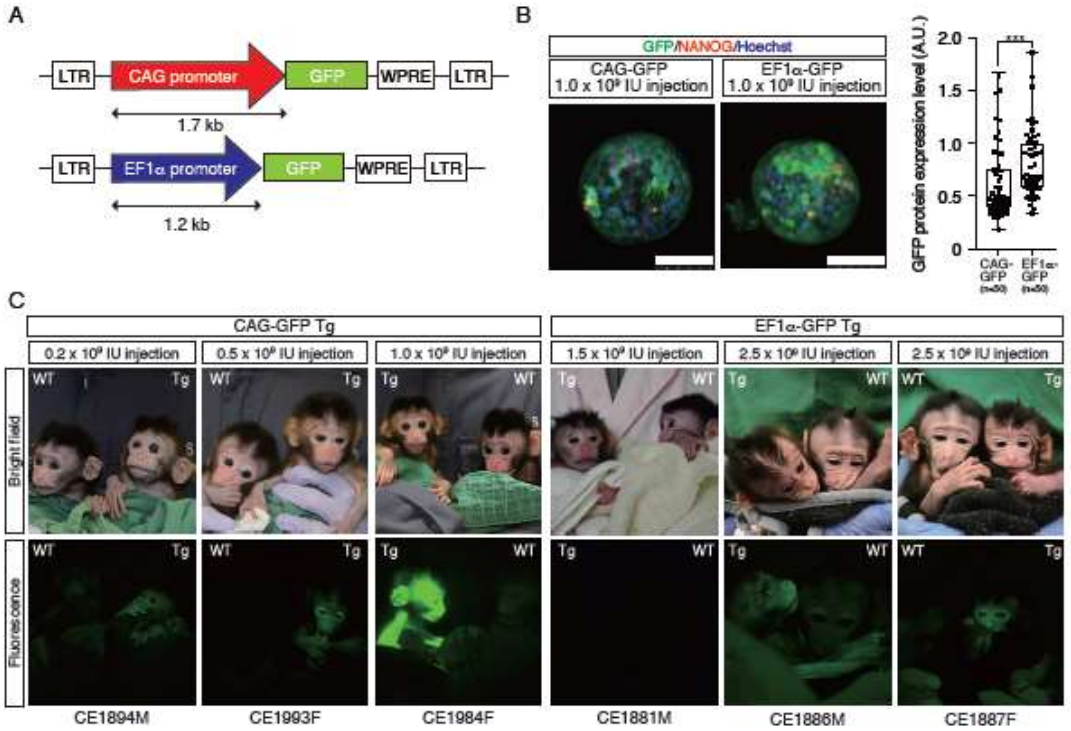

Figure 4. Generation of GFP Tg cynomolgus monkeys

A. Schematic representation of the lentiviral vector used for generation of GFP Tg monkeys. LTR: long terminal repeat. WPRE: woodchuck hepatitis posttranscriptional regulatory element. **B.** Fluorescent images of cynomolgus monkey blastocysts 7 days after infection with the lentivirus. Scale bars = 100 μm. Right panel shows Tukey box plots of GFP protein expression levels. Cell numbers are shown in brackets. A.U., arbitrary unit. Asterisks indicate statistical significance: ***$P < 0.001$. **C.** Upper panels show bright field images of the faces of CAG-GFP Tg (left) and EF1α-GFP Tg (right) offspring. Lower panels show fluorescence images of the faces of CAG-GFP Tg (left) and EF1α-GFP Tg (right) offspring.

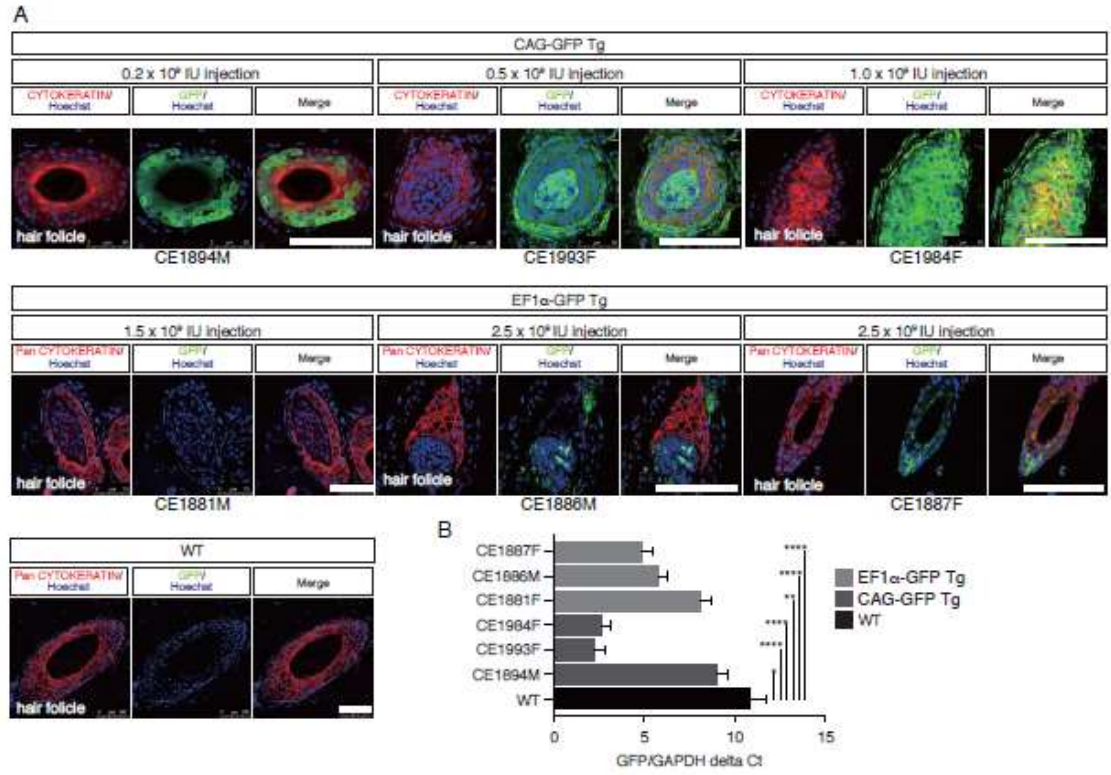

Figure 5. GFP expressions in the skins of Tg monkeys

A. Immunohistochemistry of skin tissues from CAG-GFP-Tg (upper panel) and EF1α-GFP-Tg (lower panel) offspring with anti-GFP and anti-Pan Cytokeratin antibodies detected by confocal microscopy. Images were taken under the same instrument settings (same laser intensity). Scale bars = 100 μm. **B.** GFP expression in CAG-GFP-Tg and EF1α-GFP-Tg offspring by RT–qPCR. Data are shown as the mean ± SD. Asterisks indicate statistical significance: ****$P < 0.0001$; **$P < 0.01$; *$P < 0.05$

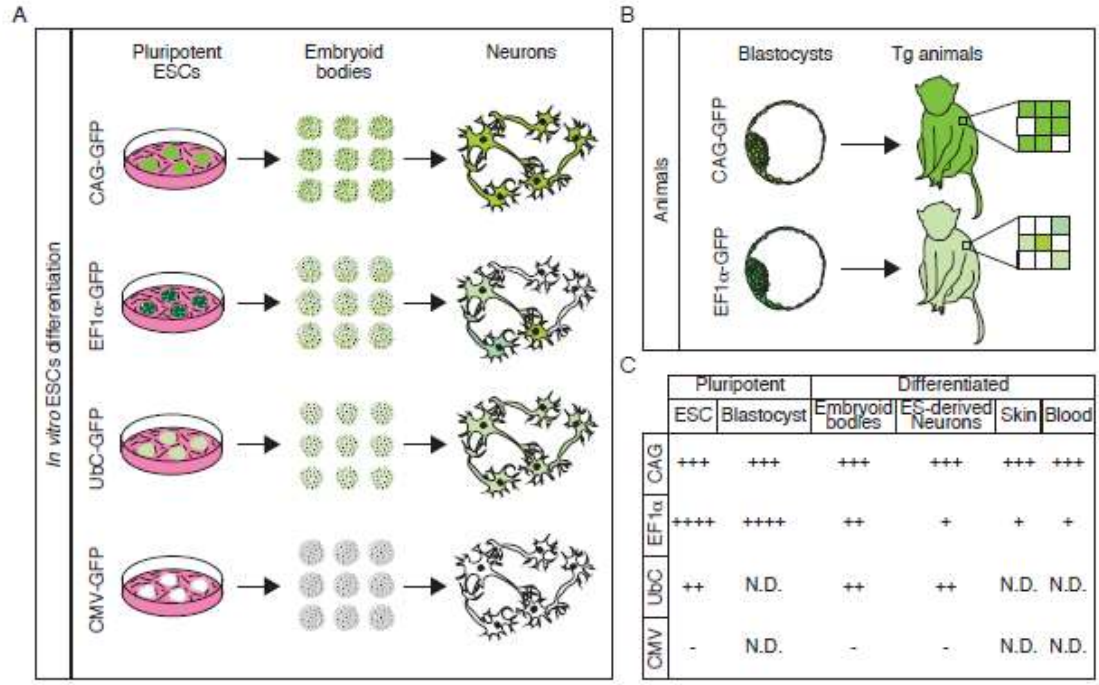

Figure 6. Summary of promoter activity measurements in transgenic cynomolgus monkeys and ESCs

A, B. Schematic of GFP expression in ESCs **(A)** and embryos **(B)**. In ESCs, the EF1α promoter drove GFP more strongly than other promoters in pluripotent ESCs, whereas the CAG promoter activity was strongest in ESC-derived tissues. However, in the Tg cynomolgus monkey, the CAG promoter drove GFP more strongly than the EF1α promoter. Notably, the EF1α promoter underwent more silencing in both ESCs and Tg monkeys. **C.** Relative promoter activity at various stages in Tg tissues. Numbers of + symbols indicate GFP expression intensity. N.D., Not determined.