Rapid and efficient stable gene transfer to mesenchymal stromal cells using a modified foamy virus vector

Nathan Paul Sweeney[1], Cathy Regan[2], Jiahui Liu[2], Antonio Galleu[3], Francesco Dazzi[3], Dirk Lindemann[4], C. Anthony Rupar[2], Myra Olga McClure[1]

[1] – Jefferiss Research Trust laboratories, Imperial College London, London, United Kingdom

[2] – Departments of Pathology and Laboratory Medicine, Biochemistry and Pediatrics, Western University, Ontario, Canada

[3] – Department of Haemato-Oncology, King's College London, London, United Kingdom

[4] - Institute of Virology, Technische Universität Dresden, Dresden, Germany

Correspondence should be addressed to MOM (mmcclure@imperial.ac.uk)

Work was carried out in London, United Kingdom and Ontario, Canada.

Professor Myra McClure

Jefferiss Research Trust laboratories, Medical School Building, Imperial College London, St. Mary's Campus, Norfolk Place, W2 1PG

Tel +44-20-7594-3902

Fax +44-20-7594-3906

Email m.mcclure@imperial.ac.uk

Short title: Efficient gene transfer to mesenchymal stromal cells

Abstract

Mesenchymal stromal cells (MSCs) hold great promise for regenerative medicine. Stable *ex vivo* gene transfer to MSCs could improve the outcome and scope of MSC therapy, but current vectors require multiple rounds of transduction, involve genotoxic viral promoters and/or the addition of cytotoxic cationic polymers in order to achieve efficient transduction. We describe a self-inactivating foamy virus vector (FVV), incorporating the simian macaque foamy virus envelope and using physiological promoters, which efficiently transduces murine MSCs (mMSCs) in a single-round. High and sustained expression of the transgene, whether GFP or the lysosomal enzyme, arylsulphatase A (ARSA), was achieved. Defining MSC characteristics (surface marker expression and differentiation potential), as well as long-term engraftment and distribution in the murine brain following intracerebroventricular delivery, are unaffected by FVV transduction. Similarly, greater than 95% of human MSCs (hMSCs) were stably transduced using the same vector, facilitating human application. This work describes the best stable gene transfer vector available for mMSCs and hMSCs.

Keywords: Gene therapy; Gene transfer; Mesenchymal stromal cells; mesenchymal stem cells; engraftment; arylsulphatase A; foamy virus; envelope.

Introduction

Mesenchymal stromal cells (MSCs) are a heterogeneous population of adult cells of mesodermal origin that can be readily isolated from bone marrow or adipose tissue and efficiently expanded *in vitro* as plastic-adherent cells. They contain a proportion of cells capable of differentiating into osteocytes, adipocytes and chondrocytes under appropriate conditions[1]. MSCs have been extensively tested in a number of clinical conditions and proved enormous potential[2]. *In vivo*, tissue resident MSCs are thought to be recruited to sites of inflammation where they play a regulatory role in wound-healing, immune modulation, angiogenesis and tissue homeostasis[3,4]. Inflammation-directed trafficking may still be retained by exogenously administered MSCs, thus making the case for their use as cancer-targeting cells[5,6] and site-directed immunosuppressive and pro-repair effector cells[7], the latter being extensively exploited for a wide-range of diseases, such as rheumatoid arthritis, type-1 diabetes, and neurological diseases such as amyotrophic lateral sclerosis and stroke[8,9].

Long-term engraftment of MSCs in the central nervous system (CNS)[10,11] facilitates their use as cellular vectors for neurodegenerative diseases, including Huntington's disease[12], Parkinson's disease[13] and lysosomal storage diseases[14] which are caused by a lysosomal enzyme deficiency. Lysosomal enzymes are tagged with a mannose-6-phosphate to enable their retrieval from the secretory pathway by its binding to the mannose-6-phosphate receptor[15]. This retrieval is leaky, resulting in some secretion of lysosomal enzyme[16]. Secreted enzyme is then retrieved by mannose-6-phosphate receptor at the cell-surface. This secrete-and-recapture system can be exploited in gene and cell therapies since enzyme expression in one cell can correct many others. For example, metachromatic leukodystrophy (MLD), one of the most common lysosomal storage diseases, is caused by a deficiency of arylsulphatase A (ARSA) causing storage of sulphatide in oligodendrocytes, Schwann cells and neurons leading to progressive demyelination and, in its most common form, death by 5 years of age[17]. Over-expression of ARSA in hematopoietic stem cells can reduce sulphatide storage caused by ARSA deficiency in the CNS of mice[18],

62 despite efficacy being dependent on the recruitment of microglia across the blood-brain-barrier. This
63 demonstrates that relatively few cells expressing ARSA could be sufficient to prevent disease. Since
64 MSCs can be safely delivered to humans, intravenously[19] and directly to the brain[20-22], a gene and MSC
65 therapy could complement or exceed a hematopoietic stem cell-based approach.

66 Any vector for gene transfer to MSCs should persist during cell division to allow *ex vivo* expansion.
67 Vectors based on retroviruses, which integrate the vector DNA into the host genome readily achieve this.
68 However, retroviral gene therapy trials have proven that these vectors carry a significant risk[23,24] due to
69 genotoxicity that is linked to the vectors' integration sites, the presence of strong viral enhancers and/or
70 transcriptional read-through[25-28]. Consequently, modern clinically relevant gene therapy vectors are
71 deleted of promoter/enhancer activity from the viral LTRs (termed self-inactivating (SIN) vectors[29]) and
72 employ non-viral physiological promoters to drive transgene expression[30]. Nonetheless, the commonly
73 used integrating vectors (gamma-retroviral or lentiviral) have so far failed to transduce MSCs efficiently
74 while retaining the aforementioned safety features[31]. Thus, a SIN vector with a physiological promoter
75 that can efficiently transduce rodent and hMSCs would boost combined gene and MSC therapies.

76 Foamy viruses are non-human, apathogenic viruses that form a distinct subgroup of the *Retroviridae*[32,33].
77 Self-inactivating foamy virus vectors (FVV) based on the prototype foamy virus (PFV) are well
78 characterised[34-38] and effective in large animal models of disease[39,40]. An historical disadvantage of FVVs
79 is that they induced a marked cytopathic effect (CPE) at a high multiplicity of infection (MOI) *in vitro*
80 due to the fusogenic nature of the PFV envelope (Env)[32]. The simian macaque foamy virus envelope
81 (SFV_{mac} Env) has recently been shown to be less fusogenic than that of PFV[41], but has not been described
82 in gene transfer.

83 This paper describes optimisation of FVVs for high transduction efficiency and transgene expression in
84 mMSCs and hMSCs. By employing the SFV_{mac} Env, high transduction efficiencies (>95%) in MSCs are

achieved from a single-round of transduction by FVV containing the cellular phosphoglycerate kinase (PGK) promoter. Viral promoters or toxic chemicals are not involved and transgene expression is high and stable for at least 10 passages post-transduction. MSC differentiation potential and surface marker expression is preserved after FVV transduction at high MOI, as is the distribution and long-term engraftment of mMSCs delivered directly to the murine brain. Thus, we describe the best existing vector for stable MSC gene transfer.

Results

The PFV Env induces syncytia formation in target cells at high MOI. Since the less fusogenic SFV_{mac} Env has not yet been tested in gene transfer, we compared the transduction efficiency in mMSCs of FVV with either the PFV or SFV_{mac} Env at different MOIs and assessed the CPE microscopically. Aside from the alternative envelopes, both vectors were identical, each carrying the same PGK-GFP construct (all FVV constructs are shown in **Fig. 1**). Both envelopes produced good vector titres, typically ranging between 10^8 and 10^9 HT1080 transducing units per ml after 100-fold concentration. Transduction efficiency was determined by flow cytometry as the percent of GFP expressing mMSCs following a single passage post-transduction (**Fig. 2a)**. At low MOIs of 1 and 5, the transduction efficiency for FVV with either the PFV or SFV_{mac} Env was similar. At higher MOIs of 10 or more, the FVV with SFV_{mac} Env resulted in significantly higher transduction efficiency than with PFV Env. This difference correlated with the MOI at which PFV Env induced extensive CPE, although syncytia formed at all MOIs tested using this envelope. In comparison, no CPE was observed even at the highest MOI of 50 for FVV with SFV_{mac} Env. The highest transduction efficiency by FVV with PFV Env was achieved between MOIs of 30 and 50 with 72% and 74% GFP-expressing cells, respectively, but was accompanied by a marked CPE. At the same MOIs, FVV with SFV_{mac} Env achieved 92 and 95% transduction efficiency, respectively, with no CPE. Representative photomicrographs of mMSCs transduced at MOI 30 with each vector are shown in

Fig. 2 b-c, demonstrating the widespread syncytia induced only by PFV Env 20 hours after vector addition. All subsequent experiments employed FVV enveloped with SFV_{mac} Env.

The phosphoglycerate kinase (PGK) and elongation factor 1α short (EFS) promoters are constitutive cellular promoters that, in contrast to viral promoters, have performed well in sensitive genotoxicity assays that measure neighbouring gene activation[30]. The ability of these promoters to achieve high transduction efficiency and maintain expression of GFP in FVV transduced mMSCs through cell expansion, a pre-requisite for cell therapy manufacturing, was compared. Vectors FVV:PGK-GFP and FVV:EFS-GFP (**Fig. 1**), were used to transduce mMSCs at MOIs of 1, 30 and 50. Transduced mMSCs were analysed by flow cytometry to determine the percent of GFP expressing cells (**Fig. 3a**) and their median fluorescence intensity (MFI) (**Fig. 3b**) after each passage post-transduction up to the 10^{th} passage (cells were passaged once confluence reached over 90% by reseeding one-tenth of the cells). This represents continuous culture over a period of 6 weeks. Representative photomicrographs of GFP fluorescence from MOI 30 transduced mMSCs is shown in **Figs. 3 c,d**. Since a MOI of 50 produced similar results to a MOI of 30 for both constructs at all passages, for clarity only data derived from a MOI of 30 are shown (**Fig. 3).**

At a MOI of both 1 and 30, FVV:PGK-GFP transduction results in a higher percentage of mMSCs expressing GFP than FVV:EFS-GFP (**Fig. 3a)**. Over 96% of mMSCs expressed GFP from 2 passages post-transduction with FVV:PGK-GFP and was sustained through subsequent passages. Comparatively, transduction with FVV:EFS-GFP at a MOI of 30 (or 50) only resulted in ~75% of mMSCs expressing GFP. At the low MOI of 1, most transduced cells contain a single vector copy (compared to multiple copies at high MOI), allowing for better analysis of expression persistence post-expansion. At this MOI, transduction with FVV:PGK-GFP enabled GFP expression in ~45% of mMSCs, stable over the 10 passages. Conversely, the 40% of mMSCs expressing GFP at 1 passage post-transduction with FVV:EFS-

GFP reduced to less than 15% by passage 4 post-transduction. Both EFS and PGK offer stable expression levels in the mMSCs that continue to express GFP, since the MFI does not change after repeated passaging (**Fig. 3b**). The PGK promoter drives approximately 5-fold higher GFP expression levels than EFS when mMSCs are transduced at a MOI of 30 from the second passage post-transduction, whereas the promoters performed similarly at one passage post-transduction. Together, **Figs. 3a and b** demonstrate the superiority of PGK as a promoter compared to EFS for FVV-mediated expression of GFP in mMSCs, providing higher expression levels, higher transduction rates and long-term stability.

A codon optimised arylsulphatase A (ARSA) open reading frame replaced GFP in our FVVs to produce FVV:PGK-ARSA and FVV:EFS-ARSA (**Fig. 1**). For high transduction efficiency, mMSCs were transduced at a MOI of 30 with these FVVs and the transduced cells collected after 1 passage. One tenth of the cells were reseeded until the 5^{th} passage post-transduction. The remaining cells were lysed to determine the intracellular ARSA activity by the ARSA assay[42] (**Fig. 4a)**. Low basal activity was detected by mMSCs transduced with FVV:PGK-GFP (control lysate). Both FVV:PGK-ARSA and FVV:EFS-ARSA induced strong ARSA activity, the highest being at passage 2 post-transduction from both vectors. The PGK promoter resulted in a 2-fold higher enzyme activity at this passage. The ARSA activity remained stable over subsequent passages for FVV:PGK-ARSA transduced mMSCs, whereas the activity in FVV:EFS-ARSA transduced mMSCs reduced between each passage with a statistically significant reduction between passages 4 and 5 post-transduction. These data are in line with those generated for GFP expression.

Since any MSC-based therapy for a lysosomal storage disease would depend on sufficient enzyme secretion and its correct processing with a mannose-6-phosphate to allow its recapture by endogenous cells, the ARSA activity in mMSC cell-culture medium and its ability to be taken up and used by MLD patients' fibroblasts was determined. The control cell-culture medium from mMSCs transduced with FVV:PGK-GFP had a low ARSA activity (hydrolysing 3.3 nmol of substrate per hour per ml) **(Fig. 4b)**,

whereas FVV:PGK-ARSA transduced mMSCs hydrolysed over 80 nmol of substrate per hour per ml, which was twofold higher than FVV:EFS-ARSA transduced mMSCs. Next, the FVV:PGK-ARSA transduced mMSCs cell-culture medium, or that from FVV:PGK-GFP (control medium), was incubated with normal (functional ARSA) or MLD patients' (ARSA deficient) fibroblasts that had been preloaded with fluorescently labelled substrate (BODIPY-sulphatide). Media from both cultures were pre-diluted to the same extent, such that 0.5 units of ARSA (the amount needed to process 0.5 pmol of substrate per hour in the ARSA assay) was added to fibroblasts in the FVV:PGK-ARSA transduced mMSC medium. This dilution caused the contribution of endogenous ARSA to be negligible. Fibroblasts from normal donors stored only low amounts of BODIPY-sulphatide in the presence of control or FVV:PGK-ARSA transduced mMSC medium, as expected (**Fig. 4c)**. Comparatively, fibroblasts from both MLD patients stored BODIPY-sulphatide in the presence of control medium. Storage was reduced when medium from FVV:PGK-ARSA transduced mMSCs was added, demonstrating that FVV:PGK-ARSA encoded ARSA is correctly processed by transduced mMSCs and can correct enzyme-deficient cells.

The effects of FVV transduction on mMSC identity and function were examined by comparing untransduced mMSCs and mMSCs transduced at MOI 30 with FVV:PGK-GFP or FVV:PGK-ARSA. A panel of antibodies targeting surface markers known to be expressed or not in mMSCs was employed and staining assessed by flow cytometry. Transduced and untransduced mMSCs stained correctly for all markers with no discernible difference in staining intensity when quantified using the Flowjo Chi squared comparison (**Fig. 5a,b**). Similarly, transduced mMSCs were able to differentiate into osteocytes, chondrocytes and adipocytes to a similar extent as untransduced mMSCs when cultured under appropriate conditions (**Fig. 5c-k**).

Given the prospect of using MSCs for therapy of diseases affecting the CNS, a sensitive qPCR targeting the Y-chromosome to allow detection of male mMSCs delivered directly to the brains of female mice[43] was established. To determine if FVV transduction affected the long-term engraftment capability of

mMSCs, untransduced or FVV:EFS-ARSA transduced male mMSCs were each injected into the right lateral ventricle of 6 female mice. Three months post-injection, treated mice were sacrificed and their brains crudely sectioned into 8 blocks (**Fig. 6a**). Genomic DNA was isolated from each block and the amount of male (mMSC-derived) DNA was determined by qPCR. The number of male genomes present per million total genomes (male and female) is shown for sections 1-8 in **Fig. 6b**. Data points are only shown for sections that exceeded the detection limit of approximately 10 male genomes per million. Section 6, which includes the injected ventricle (**Fig. 6a**), was the most likely section to contain detectable levels of mMSC DNA with 11 of the 12 injected mice having detectable levels. This section also tended to contain a higher proportion of male DNA than other sections. Section 3, containing the non-injected lateral ventricle, also featured high levels of male DNA in most treated mice. The cerebellum (section 8) was the least likely section to harbour mMSC DNA with only 2 mice having sufficient numbers for detection, both of which had been injected with FVV transduced mMSCs. Higher numbers of mice had detectable levels of male DNA in all other sections, showing that mMSCs migrate from the injected lateral ventricle throughout the brain. There was no discernible difference in distribution or level of engraftment between transduced and untransduced MSCs.

To test whether FVV transduced mMSCs could maintain transgene expression *in vivo*, 6 adult mice were injected in their right lateral ventricle with FVV:PGK-GFP transduced mMSCs. Half the mice were sacrificed immediately post-injection and the other half after 45 days. Coronal cryosections of their brains were examined for direct GFP fluorescence (**Fig. 6 c-f**). Evidence of GFP expression was found for both time-points. The injected lateral ventricle contained many GFP-expressing cells immediately post-injection and was enlarged, while the non-injected ventricle also contained some GFP-expressing cells. After 45-days, GFP-expressing cells were found predominantly, but not exclusively, associated with the choroid plexus; along the needle track route in the parenchyma; and in the glomerular layer of the

olfactory bulb. This shows that the PGK promoter remains active in FVV transduced mMSCs that have grafted long-term in the murine brain.

To determine whether a FVV-based MSC therapy could be translated from pre-clinical work in mice to clinical use in humans, the GFP-encoding FVVs were tested on hMSCs obtained from 3 different donors. At an early passage number (2-4), the hMSCs were transduced with FVV:PGK-GFP or FVV:EFS-GFP at different MOIs. At a MOI of 50, we also tested transduction of these vectors employing the PFV Env. The results (**Fig. 7a)** show that, in contrast to results in mMSCs, the PFV Env achieved similar transduction efficiencies to SFV_{mac} Env at high MOI. However, CPE was again induced by PFV Env, although to a lesser extent than in mMSCs at MOI 50 (not shown). At all MOIs and with either envelope, FVV:PGK-GFP and FVV:EFS-GFP perform similarly. For SFV_{mac} Env containing FVVs, under 10% of hMSCs expressed GFP at MOI 1. Each increase in MOI tested resulted in a higher percent of GFP expressing hMSCs. At a MOI of 100, the highest tried, approximately 95% of hMSCs expressed GFP. Normal morphology and no CPE was observed by microscopy at this high MOI (**Fig. 7 b-c**). The average time between passages (one-fifth cells reseeded at each passage) for transduced and untransduced hMSCs was 4 days between each passage until the 7^{th} post-transduction, where the time taken between passages doubled as cells putatively entered senescence. Transduced hMSCs retained normal osteogenic differentiation potential when tested at 6 passages post-transduction (not shown). Therefore, FVV transduction did not affect proliferation or function of hMSCs when assessed by these measures.

Flow cytometry was performed for hMSCs transduced at MOIs of 1 and 100 with both FVV:PGK-GFP and FVV:EFS-GFP following each passage, until the 10^{th} post-transduction, to monitor the percent of GFP expressing cells and their MFI (**Fig. 7 d and e**, respectively). During cell expansion, until growth slowed, neither the percent of hMSCs expressing GFP nor their MFI changed for hMSCs transduced with FVV:PGK-GFP or FVV:EFS-GFP at MOI 1 or 100. Both promoters exerted similar activity. At low MOI, there was variation in transduction efficiency of hMSCs from different donors, but this became

unapparent at high MOI. Following the putative entering of senescence, the percent of hMSCs expressing GFP reduced at a MOI of 1 from ~7% pre-senescence to less than 1% at the 9th passage post-transduction with FVV:EFS-GFP. A reduction was also observed for hMSCs transduced at low MOI with FVV:PGK-GFP. At high MOI, the percent of GFP-expressing cells in FVV:PGK-GFP transduced hMSCs were unaffected by putative senescence, although the MFI became more variable between hMSCs from different donors. For just 1 of the 3 hMSCs transduced with FVV:EFS-GFP at high MOI, the percentage expressing GFP dropped from over 95% pre-senescence to 50.1% at the 8th passage post-transduction, resulting in the reduction in the average GFP expressing cells transduced with this vector. These data show that, at least pre-senescence, either of the tested FVVs are highly suited for high and stable transgene expression.

Discussion

All previous reports on FVVs have used the PFV envelope despite its toxicity to cells at high MOI. The SFV_{mac} Env is less toxic since, in contrast to PFV Env, it only has significant fusion activity at low pH[41], thus preventing fusion at the cell membrane. The PFV Env has extremely broad tropism, with only a single zebrafish cell line (Pac2) found to be resistant to PFV infection[44]. If the SFV_{mac} Env permits broad tropism, its lack of toxicity would be advantageous. We have shown that for mMSCs, the SFV_{mac} Env not only allows FVV use at high MOI without inducing syncytia, but unexpectedly enables higher transduction efficiencies to be achieved. Making this simple change may also benefit FVV transduction of other cell types.

We compared the activity of two constitutive cellular promoters, EFS and PGK, in mMSCs and hMSCs over 10 passages. Unexpectedly, the promoter affected the observed transduction efficiency in mMSCs.

Since our measure of transduction efficiency depends on detectable GFP expression in each transduced cell, this result is likely caused by the absence of expression in some (FVV:EFS-GFP) transduced cells. This may be due to a combination of the unique integration site selection and/or the inherent heterogeneity within MSC populations. By monitoring expression through cell expansion (Fig. 3), it becomes apparent that the EFS promoter is subjected to putative silencing, as the percent of GFP-expressing cells decreased during expansion. In contrast, the PGK promoter conferred stable expression which peaked at 2 passages post-transduction. This late peak may be explained by the FVV's dependence on cell division for integration and transgene expression[45]. The silencing of EFS may also have played a role in reducing transduction rates using this promoter by counteracting the increase in GFP-expressing cells, since transduction of mMSCs with FVV:EFS-GFP at low MOI did not show a similar peak. Interestingly, neither the expression levels nor the observed transduction rates were influenced in hMSCs by promoter choice (Fig. 7), indicating different activities of these housekeeping genes between the species of MSCs tested.

Using the lysosomal storage diseases as a proposed target, we showed that ARSA overexpression by FVVs in mMSCs results in strong enzyme activity with a significant amount being secreted. This secreted enzyme was appropriately processed since it could be used by ARSA deficient cells to clear stored substrate. Furthermore, we have shown that mMSC engraftment in the CNS is not affected by FVV transduction and *in vivo* transgene expression is maintained for at least 45 days. Our long-term engraftment levels and distribution was consistent with published results by an independent group using unmodified mMSCs[43]. Lack of a specific marker to identify MSCs *in vivo* prevents us from determining whether the GFP-expressing (MSC-derived) cells have maintained MSC identity or not. However, for the proposed use of long-term transgene expression, long-term survival and transgene expression is more important than their final identity. Thus, the long-term engraftment and FVV-mediated transgene

expression *in vivo* that we demonstrate underpins proposals to use an MSC-based gene therapy approach for the treatment of lysosomal storage diseases affecting the CNS.

Our work has achieved both high and stable transduction efficiency in both mMSCs and hMSCs, so the same vector can be employed for pre-clinical and clinical applications. Importantly, FVV transduction is not enhanced by polybrene[46], shown to affect hMSC proliferation capacity[47]. No additives were used to achieve high transduction efficiency with FVV. Although high transduction efficiency in hMSCs has been reported for lentiviral vectors[31,48-50], these included viral promoters/enhancers. Even with a viral promoter driving transgene expression, a SIN lentiviral vector required 3 rounds of transduction to get 92% efficiency in hMSCs[50]. Given that viral promoters in a retroviral context have been strongly linked to oncogenesis in other cell types[27,30,51], these vector designs are unlikely to be approved for clinical use. The vector we describe in this paper does not use viral promoters and, thus, has stronger clinical prospects.

In addition to our vector being devoid of viral promoters, FVVs have innate properties that make them favourable for gene therapy, including a potent transcriptional terminator that prevents transcriptional read-through[52], an integration site bias that does not favour active genes or their regulatory regions[53,54] and being derived from a non-human apathogenic virus[55]. Unfortunately, directly testing our vector's safety in MSCs is challenging. No transformation assays exist for hMSCs which senesce *in vitro* following long-term culture, while mMSCs undergo a pre-transformation stage when cultured at ambient oxygen levels *in vitro*, making them unsuitable for such analysis. However, we monitored GFP expression during significant *in vitro* expansion yet saw no indication of clonal dominance, since the MFI and the percentage of GFP-expressing cells were stable in FVV:PGK-GFP transduced cells (Figs. 3 and 7). Moreover, hMSCs became senescent at the same passage number in transduced and untransduced cells, indicating that a sub-population was not transformed. Nevertheless, without sensitive transformation assays available for MSCs, we are unable to draw conclusions on the relative safety of our

vector in MSCs. Integration site analysis and testing for oncogenic potential in NOD SCID mice may be appropriate experiments to be carried out prior to clinical translation.

In conclusion, we have developed a potentially safe integrating vector that is highly efficient in both mouse and human MSCs. For both species of MSC, over 95% express transgene stably for at least 10 passages after transduction. All future combined gene and MSC therapies should take advantage of the FVV described in this work.

Materials and Methods

Cell isolation and culture

All cells were cultured under sterile conditions at 5% CO_2 and 37℃ in a humidified incubator. The adherent human embryonic kidney and human fibrosarcoma cell lines; HEK-293T[56] and HT1080[57]; were cultured in DMEM containing 10% FBS. mMSCs were isolated from the bone marrow of 4-6 week old male C57BL/6 mice. Bone marrow was flushed from the femurs and tibias of 3 mice, pooled and cultured using the Mesencult Mouse Proliferation kit by StemCell Technologies, UK according to the manufacturer's recommended protocol with the exception that dissociation was performed using TrypLE Express (ThermoFisher Scientific, UK).

Clinical grade human bone marrow MSCs were produced in accordance of the Regulation (EC) No 1394/2007 of the European Parliament and of the Council on advanced therapy medicinal products and amending Directive 2001/83/EC and Regulation (EC) No 726/2004. For isolation, 2 ml of bone marrow aspirate was collected from the iliac crests of healthy donors into 100 µl preservative-free heparin. Within 24 hours, cells were seeded at a density of 15-40 000 cells per cm^2 in αMEM containing antibiotics and

10% FBS. After 3 days, non-adherent cells were discarded and adherent cells cultured until confluence. Cells were maintained using αMEM containing 5% human platelet lysate (Stemulate PL-NH from Cook Medical, UK). When cells reached over 90% confluence they were dissociated using TrypLE express and reseeded at a density of 4000 per cm^2. All hMSCs were tested for surface marker expression by flow cytometry and satisfied the recommended minimal criteria[58]. Over 95% were positive for CD105, CD73 and CD90 expression while less than 2% were positive for the negative markers CD45, CD34, CD3, CD14, CD19 and HLA-DR.

FVV production

Generation of transfer vector constructs is described in the supplementary material. FVVs were produced in HEK-293T cells transfected with a 4-plasmid system using PEImax (Polysciences, Germany) at a ratio of 3:1 PEI:DNA. The 4-plasmid system comprised of one of the pDΦ- transfer plasmids, pcoPG4 (encoding PFV Gag), pcoPPwt (encoding PFV pol), and either pcoPE or pcoSE (encoding PFV or SFV_{mac} Env, respectively) in the ratio 52:13:6:4, respectively. Plasmids pcoPG4, pcoPPwt and pcoPE have been described previously[59]. In a typical transfection, 5 x 10^6 HEK-293T cells were seeded per 55 cm^2 round culture dish. The next day, cells were transfected with 15 μg of DNA. Following transfection, published protocols for FVV collection, concentration and storage[60] were followed.

FVV transduction and titration

For HT1080 cell transduction, 10^4 cells were seeded per cm^2 surface area. After 16-24 hours, FVV was added. After a further 16-24 hours, medium was replaced. For vectors expressing GFP, cells were collected at confluence and the percent of GFP expressing cells (using a vector dilution that gave between 1 and 15% GFP expressing cells) was determined by flow cytometry. The titre was determined by multiplying the proportion of GFP expressing cells by the number of cells at the time of vector addition. For ARSA encoding vectors, qPCR analysis determined the FVV DNA content 1 passage post-transduction in transduced HT1080 cells relative to that of cells transduced with GFP vector of known

titre. For this, the primers 203-F (AGATTGTACGGGAGCTCTTCAC), 203-R (CAGAAAGCATTGCAATCACC) and dual-labelled probe 203-P (FAM-TACTCGCTGCGTCGAGAGTGTACGA-BHQ-1), which target the FVV LTR, were employed. FVV DNA content was normalised to the albumin gene using published primers and probe[61].

To transduce mMSCs, 2500 cells were seeded per cm^2 in the presence of FVV. The cells and vector were centrifuged at 1200g for 90 minutes at 30℃ then cultured normally. After 16-24 hours, medium was replaced. For hMSCs, 4000 cells were seeded per cm^2. After 16-24 hours, FVV was added and the cultures centrifuged as for mMSCs. Medium was replaced after 5-8 hours.

Flow cytometry

To assess the percentage of GFP expressing cells and their MFI, at least 10 000 single cells were acquired using a Beckton Dickinson LSRII. Single cells were selected using forward and side scatter parameters. Untransduced cells served as a negative control to set the gates in the 488-530/30 channel to determine the percentage of GFP expressing cells. For mMSCs, which exhibited strong autofluorescence, the GFP signal in the 488-530/30 channel was plotted against the signal in the 488-610/20 channel. This strategy distinguished between GFP (stronger in 488-530/30 than 488-610/20) and autofluorescence (similar in both channels). Example plots are shown in Supplementary Fig. 1. The MFI was calculated by subtracting the median signal intensity of the GFP negative population from that of the GFP positive population. Surface marker expression analysis for mMSCs was carried out using the Mouse Mesenchymal Stem Cell Marker Antibody Panel (R&D systems, UK) according to the manufacturer's recommended protocol. Data was acquired using a Beckton Dickinson LSRFortessa. Data analysis was performed using FlowJo.

Functional assays

To quantify intracellular ARSA activity, cells were collected 5 days after reaching confluence and lysed on ice in 20 mM Tris-HCl, pH 8.0, 137 mM NaCl, 1% Triton X-100 and 2 mM EDTA. Lysates were

362 cleared by centrifugation at 16 000 g and supernatants collected. Protein concentration was determined
363 using the DC protein assay (Bio-Rad, UK). The ARSA assay performed as previously described[42] using
364 between 1 and 2 µg protein per reaction. To determine ARSA activity in cell-culture medium, mMSCs (2
365 passages post-transduction) were grown to confluence and complete medium change was performed.
366 After 5 days, the medium was collected and filtered through a 0.45 µm cellulose acetate syringe filter and
367 stored at -20℃ until use. For each reaction, 40 µl medium was added. All reactions were performed in
368 triplicate or quadruplicate in a 96-well microtiter plate.

369 BODIPY-sulphatide was produced using Lysosulphatide (Matreya, USA), BODIPY FL C16 (Life
370 Technologies, Canada) and dicyclohexylcarbodiimide (Sigma-Aldrich, Canada) as previously described[62].
371 Patient fibroblasts (MLD or healthy controls) were grown to 75% confluence and BODIPY-sulphatide
372 was added to a final concentration of 6.7 mM. After 24 hours, cells were rinsed twice with PBS and
373 diluted MSC cell-culture supernatant added. Cells were collected for high-performance liquid
374 chromatography analysis after a further 24 hours. BODIPY-lactosylceramide (synthesised as described
375 for BODIPY-sulphatide) was added as a recovery standard and lipids were extracted using chloroform
376 and methanol as described[63]. Chromatographic separations were performed using a Luna C18 column
377 (Phenomenex; Canada). The mobile phase consisted of solvent A: methanol:water (1:1; v/v) and solvent
378 B: tetrahydrofuran:methanol (4:1; v/v). The flow rate was 1 ml/min. When eluting the column, the mobile
379 phase was increased from 40% solvent B to 100% solvent B in 20 minutes, held at 100% solvent B for 10
380 min, then decreased back to 40% solvent B and held for 10 minutes. The fluorescence detector was set
381 with an excitation wavelength of 502 nm and emission wavelength of 530 nm.

382 The differentiation potential of mMSCs into osteocytes, chondrocytes and adipocytes was tested using
383 StemPro Osteogenesis Differentiation Kit, StemPro Chondrogenesis Kit (both ThermoFisher Scientific)
384 or MesenCult Adipogenic Stimulatory Supplements, mouse (Stemcell technologies), respectively. For

hMSCs, osteogenic differentiation was tested as for mMSCs. Manufacturers' protocols were followed in all instances. Differentiation was confirmed by staining with Alizarin Red, Alcian Blue or Oil Red O (all from Sigma-Aldrich, UK), respectively.

Stereotaxic injections and tissue processing

$ARSA^{-/-}$ mice, which contain a large deletion in the ARSA gene (Hess et al., 1996), had previously been bred onto a C57BL/6 background at the University of Western Ontario, Canada. Procedures were carried out in compliance with the guidelines set by the Canadian Council for Animal Care. Stereotaxic injections were carried out as previously described[64]. Following euthanasia, brains were collected and submerged in RNAlater (Qiagen, Canada), then frozen at -80℃. Brains were either sectioned using a cryostat for microscopy or cut into blocks for genomic DNA extraction. Genomic DNA was extracted using the QIAamp DNA mini kit (Qiagen, UK). The amount of male DNA present in female brains was determined by qPCR according to published protocols[43]. The standard curve was generated by adding genomic DNA from mMSCs to genomic DNA extracted from female $ARSA^{-/-}$ mouse brains.

Microscopy

Photomicrographs of tissue sections were acquired using with the Openlab imaging software (Perkin Elmer, Canada) connected to an inverted fluorescence Leica DM IRB microscope (Leica Microsystems, Canada). Fluorescence and light microscopy of cells *in vitro* was performed using a Nikon Eclipse TE-2000s and images were captured using the Nikon ACT-1 software (Nikon, UK).

Statistical analyses

Graphing and statistical analyses were carried out using GraphPad Prism version 6.07. Asterisks denote the P value from statistical tests (detailed in Figure legends) where $P<0.05$=*, $P<0.01$=**, $P<0.001$=***.

Acknowledgments

This work was funded by a Wellcome trust PhD studentship 093610/Z/10/A awarded to NPS and a grant from Bethany's Hope Foundation (CAR). The authors declare no conflict of interest.

Supplementary Materials and Methods

Construction of FVV transfer plasmids

The FVV transfer plasmids were modified from pDΦ[34], a gift from D. Russell. The murine PGK promoter was PCR amplified from pQ-PGK-FLAG-puro (a gift from G. Maertens) using primers PGK-F (GCATCGATTTCTACCGGGTAGGGGAGGC) and PGK-R (GCGGTACCAGGTCGAAAGGCCCGGAGATG). The EFS promoter was PCR amplified from the EF1α promoter in pWPT-GFP (a gift from J. Luban) using primers EFS-F GCATCGATTGGCTCCGGTGCCCGTCAGT and EFS-R (GCGGTACCCGCGTCACGACACCTGTGTT). Both promoters were inserted between the ClaI-KpnI restriction sites of pDΦ. The enhanced GFP open-reading frame was PCR amplified from pWPT-GFP using primers GFP-F (GCGGTACCATGGTGAGCAAGGGCGAGGA) and GFP-R2 (GCGCGGCCGCAAGCTTCTAGCTACTAGCTAGTCGAG) and cloned into the PCR4-TOPO vector using the Zero Blunt TOPO PCR Cloning Kit for Sequencing (ThermoFisher scientific). The woodchuck hepatitis virus post-transcriptional regulatory element (PRE) was PCR amplified from pLVx-EF1α-IRES-mCherry (Clontech, UK) using primers WPRE-F (GCAAGCTTAATCAACCTCTGGATTACAA) and WPRE-R (GCGCGGCCGCCAGGCGGGGAGGCGGCCCAA) then inserted into PCR4-TOPO containing GFP between HindIII and NotI restriction sites. GFP-PRE was added between KpnI-NotI restriction sites to generate plasmids pDΦ-PGK-GFP-wPRE and pDΦ-EFS-GFP-wPRE. A codon optimised ARSA open-reading frame was amplified from pJO4-ASA (DNA2.0, Menlo Park, CA) using

primers ARSA-F (GCATCGATGGTACCATGGGTGCGCCCAGATCGTT) and ARSA-R (GCGGATCCTCACGCATGCGGGTCCGGAC) and inserted between the KpnI and BamHI restriction sites of pDΦ containing EFS or PGK. The optimised PRE[65] was amplified from pENTR-L5-oPRE-L2 (Addgene plasmid 32414) using primers oPRE-F (GCGGATCCTATACAAAAGTTGTGGAGCA) and oPRE-R (CAGCGGCCGCACGACAACACCACGGAAT) then inserted between BamHI and NotI restriction sites to generate pDΦ-PGK-ARSA-oPRE and pDΦ-EFS-ARSA-oPRE. Sequencing confirmed the correct insert for all plasmids.

References

1 Pittenger, M. F. *et al.* Multilineage potential of adult human mesenchymal stem cells. *Science* **284**, 143-147 (1999).

2 Dazzi, F. & Horwood, N. J. Potential of mesenchymal stem cell therapy. *Current opinion in oncology* **19**, 650-655, doi:10.1097/CCO.0b013e3282f0e116 (2007).

3 Yagi, H. *et al.* Mesenchymal stem cells: Mechanisms of immunomodulation and homing. *Cell Transplant* **19**, 667-679, doi:10.3727/096368910x508762 (2010).

4 Wu, Y., Chen, L., Scott, P. G. & Tredget, E. E. Mesenchymal stem cells enhance wound healing through differentiation and angiogenesis. *Stem Cells* **25**, 2648-2659, doi:10.1634/stemcells.2007-0226 (2007).

5 Shah, K. Mesenchymal stem cells engineered for cancer therapy. *Advanced drug delivery reviews* **64**, 739-748, doi:10.1016/j.addr.2011.06.010 (2012).

6 Hu, Y. L., Fu, Y. H., Tabata, Y. & Gao, J. Q. Mesenchymal stem cells: a promising targeted-delivery vehicle in cancer gene therapy. *Journal of controlled release : official journal of the Controlled Release Society* **147**, 154-162, doi:10.1016/j.jconrel.2010.05.015 (2010).

7 Prockop, D. J. "Stemness" Does Not Explain the Repair of Many Tissues by Mesenchymal Stem/Multipotent Stromal Cells (MSCs). *Clinical Pharmacology & Therapeutics* **82**, 241-243, doi:10.1038/sj.clpt.6100313 (2007).

8 Meirelles Lda, S., Fontes, A. M., Covas, D. T. & Caplan, A. I. Mechanisms involved in the therapeutic properties of mesenchymal stem cells. *Cytokine & growth factor reviews* **20**, 419-427, doi:10.1016/j.cytogfr.2009.10.002 (2009).

9 Uccelli, A., Moretta, L. & Pistoia, V. Mesenchymal stem cells in health and disease. *Nat Rev Immunol* **8**, 726-736 (2008).

10 Phinney, D. G. *et al.* Murine mesenchymal stem cells transplanted to the central nervous system of neonatal versus adult mice exhibit distinct engraftment kinetics and express receptors that guide neuronal cell migration. *Stem Cells Dev* **15**, 437-447, doi:10.1089/scd.2006.15.437 (2006).

11 Isakova, I. A., Baker, K., Dufour, J., Gaupp, D. & Phinney, D. G. Preclinical evaluation of adult stem cell engraftment and toxicity in the CNS of rhesus macaques. *Mol Ther* **13**, 1173-1184, doi:10.1016/j.ymthe.2005.12.014 (2006).

12 Olson, S. D. *et al.* Genetically engineered mesenchymal stem cells as a proposed therapeutic for Huntington's disease. *Mol Neurobiol* **45**, 87-98, doi:10.1007/s12035-011-8219-8 (2012).

13 Kitada, M. & Dezawa, M. Parkinson's disease and mesenchymal stem cells: potential for cell-based therapy. *Parkinson's disease* **2012**, 873706, doi:10.1155/2012/873706 (2012).

14 Phinney, D. G. & Isakova, I. A. Mesenchymal stem cells as cellular vectors for pediatric neurological disorders. *Brain Res* **1573**, 92-107, doi:10.1016/j.brainres.2014.05.029 (2014).

15 Coutinho, M. F., Prata, M. J. & Alves, S. Mannose-6-phosphate pathway: A review on its role in lysosomal function and dysfunction. *Molecular Genetics and Metabolism* **105**, 542-550, doi:http://dx.doi.org/10.1016/j.ymgme.2011.12.012 (2012).

16 Kornfeld, S. Trafficking of lysosomal enzymes. *The FASEB Journal* **1**, 462-468 (1987).

17 Gieselmann, V. & Krägeloh-Mann, I. Metachromatic Leukodystrophy - An Update. *Neuropediatrics* **41**, 1,6, doi:10.1055/s-0030-1253412 (2010).

18 Biffi, A. *et al.* Gene therapy of metachromatic leukodystrophy reverses neurological damage and deficits in mice. *The Journal of clinical investigation* **116**, 3070-3082 (2006).

19 Lalu, M. M. *et al.* Safety of cell therapy with mesenchymal stromal cells (SafeCell): a systematic review and meta-analysis of clinical trials. *PLoS ONE* **7**, e47559, doi:10.1371/journal.pone.0047559 (2012).

20 Venkataramana, N. K. *et al.* Open-labeled study of unilateral autologous bone-marrow-derived mesenchymal stem cell transplantation in Parkinson's disease. *Translational Research* **155**, 62-70, doi:http://dx.doi.org/10.1016/j.trsl.2009.07.006 (2010).

21 Mazzini, L. *et al.* Mesenchymal stem cell transplantation in amyotrophic lateral sclerosis: A Phase I clinical trial. *Exp Neurol* **223**, 229-237, doi:10.1016/j.expneurol.2009.08.007 (2010).

22 Karussis, D. *et al.* Safety and immunological effects of mesenchymal stem cell transplantation in patients with multiple sclerosis and amyotrophic lateral sclerosis. *Archives of neurology* **67**, 1187-1194, doi:10.1001/archneurol.2010.248 (2010).

23 Gaspar, H. B. *et al.* Long-term persistence of a polyclonal T cell repertoire after gene therapy for X-linked severe combined immunodeficiency. *Sci Transl Med* **3**, 97ra79, doi:10.1126/scitranslmed.3002715 (2011).

24 Hacein-Bey-Abina, S. *et al.* Efficacy of gene therapy for X-linked severe combined immunodeficiency. *N Engl J Med* **363**, 355-364, doi:10.1056/NEJMoa1000164 (2010).

25 Cesana, D. *et al.* Whole transcriptome characterization of aberrant splicing events induced by lentiviral vector integrations. *The Journal of clinical investigation* **122**, 1667-1676, doi:10.1172/jci62189 (2012).

26 Williams, D. A. & Thrasher, A. J. Concise Review: Lessons Learned From Clinical Trials of Gene Therapy in Monogenic Immunodeficiency Diseases. *Stem cells translational medicine* **3**, 636-642, doi:10.5966/sctm.2013-0206 (2014).

27 Knight, S., Bokhoven, M., Collins, M. & Takeuchi, Y. Effect of the internal promoter on insertional gene activation by lentiviral vectors with an intact HIV long terminal repeat. *J Virol* **84**, 4856-4859, doi:10.1128/jvi.02476-09 (2010).

28 Bokhoven, M. *et al.* Insertional gene activation by lentiviral and gammaretroviral vectors. *J Virol* **83**, 283-294, doi:10.1128/jvi.01865-08 (2009).

29 Miyoshi, H., Blömer, U., Takahashi, M., Gage, F. H. & Verma, I. M. Development of a Self-Inactivating Lentivirus Vector. *Journal of Virology* **72**, 8150-8157 (1998).

30 Zychlinski, D. *et al.* Physiological Promoters Reduce the Genotoxic Risk of Integrating Gene Vectors. *Mol Ther* **16**, 718-725 (2008).

31 McGinley, L. *et al.* Lentiviral vector mediated modification of mesenchymal stem cells & enhanced survival in an in vitro model of ischaemia. *Stem cell research & therapy* **2**, 12, doi:10.1186/scrt53 (2011).

32 Rethwilm, A. Molecular biology of foamy viruses. *Med Microbiol Immunol* **199**, 197-207, doi:10.1007/s00430-010-0158-x (2010).

33 Khan, A. S. Simian foamy virus infection in humans: prevalence and management. *Expert review of anti-infective therapy* **7**, 569-580, doi:10.1586/eri.09.39 (2009).

34 Trobridge, G., Josephson, N., Vassilopoulos, G., Mac, J. & Russell, D. W. Improved foamy virus vectors with minimal viral sequences. *Mol Ther* **6**, 321-328 (2002).

35 Heinkelein, M. *et al.* Improved primate foamy virus vectors and packaging constructs. *J Virol* **76**, 3774-3783 (2002).

36 Hill, C. L., Bieniasz, P. D. & McClure, M. O. Properties of human foamy virus relevant to its development as a vector for gene therapy. *Journal of General Virology* **80**, 2003-2009 (1999).

37 Erlwein, O., Bieniasz, P. D. & McClure, M. O. Sequences in pol are required for transfer of human foamy virus-based vectors. *J Virol* **72**, 5510-5516 (1998).

38 Wiktorowicz, T., Peters, K., Armbruster, N., Steinert, A. F. & Rethwilm, A. Generation of an improved foamy virus vector by dissection of cis-acting sequences. *J Gen Virol* **90**, 481-487, doi:10.1099/vir.0.006312-0 (2009).

39 Bauer, T. R. *et al.* Successful treatment of canine leukocyte adhesion deficiency by foamy virus vectors. *Nat Med* **14**, 93-97, doi:http://www.nature.com/nm/journal/v14/n1/suppinfo/nm1695_S1.html (2008).

40 Bauer, T. R., Jr. *et al.* Long-Term Follow-up of Foamy Viral Vector-Mediated Gene Therapy for Canine Leukocyte Adhesion Deficiency. *Mol Ther*, doi:10.1038/mt.2013.34 (2013).

41 Stirnnagel, K. *et al.* Differential pH-dependent cellular uptake pathways among foamy viruses elucidated using dual-colored fluorescent particles. *Retrovirology* **9**, 71, doi:10.1186/1742-4690-9-71 (2012).

42 Rip, J. W. & Gordon, B. A. A Simple Spectrophotometric Enzyme Assay with Absolute Specificity for Arylsulfatase A. *Clinical biochemistry* **31**, 29-31, doi:10.1016/s0009-9120(97)00142-2 (1998).

43 McBride, C., Gaupp, D. & Phinney, D. G. Quantifying levels of transplanted murine and human mesenchymal stem cells in vivo by real-time PCR. *Cytotherapy* **5**, 7-18, doi:10.1080/14653240310000038 (2003).

44 Stirnnagel, K. *et al.* Analysis of Prototype Foamy Virus particle-host cell interaction with autofluorescent retroviral particles. *Retrovirology* **7**, 45 (2010).

45 Patton, G. S., Erlwein, O. & McClure, M. O. Cell-cycle dependence of foamy virus vectors. *J Gen Virol* **85**, 2925-2930, doi:10.1099/vir.0.80210-0 (2004).

46 Loh, P. C. & Ang, K. S. Replication of Human Syncytium-Forming Virus in Human Cells: Effect of Certain Biological Factors and Selective Chemicals. *Journal of Medical Virology* **7**, 67-73, doi:10.1002/jmv.1890070108 (1981).

47 Lin, P., Correa, D., Lin, Y. & Caplan, A. I. Polybrene Inhibits Human Mesenchymal Stem Cell Proliferation during Lentiviral Transduction. *PLoS ONE* **6**, e23891, doi:10.1371/journal.pone.0023891 (2011).

48 Zhang, X. Y. *et al.* Lentiviral vectors for sustained transgene expression in human bone marrow-derived stromal cells. *Mol Ther* **5**, 555-565, doi:10.1006/mthe.2002.0585 (2002).

49 Lin, P. *et al.* Efficient lentiviral transduction of human mesenchymal stem cells that preserves proliferation and differentiation capabilities. *Stem cells translational medicine* **1**, 886-897, doi:10.5966/sctm.2012-0086 (2012).

50 Van Damme, A. *et al.* Efficient Lentiviral Transduction and Improved Engraftment of Human Bone Marrow Mesenchymal Cells. *STEM CELLS* **24**, 896-907, doi:10.1634/stemcells.2003-0106 (2006).

51 Montini, E. *et al.* The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. *The Journal of clinical investigation* **119**, 964-975, doi:10.1172/jci37630 (2009).

52 Hendrie, P. C., Huo, Y., Stolitenko, R. B. & Russell, D. W. A Rapid and Quantitative Assay for Measuring Neighboring Gene Activation by Vector Proviruses. *Mol Ther* **16**, 534-540 (2008).

53 Trobridge, G. D. *et al.* Foamy virus vector integration sites in normal human cells. *Proc Natl Acad Sci U S A* **103**, 1498-1503, doi:10.1073/pnas.0510046103 (2006).

54 Nowrouzi, A. *et al.* Genome-wide mapping of foamy virus vector integrations into a human cell line. *Journal of General Virology* **87**, 1339-1347, doi:10.1099/vir.0.81554-0 (2006).

55 Betsem, E., Rua, R., Tortevoye, P., Froment, A. & Gessain, A. Frequent and recent human acquisition of simian foamy viruses through apes' bites in central Africa. *PLoS Pathog* **7**, e1002306, doi:10.1371/journal.ppat.1002306 (2011).

56 DuBridge, R. B. *et al.* Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Molecular and cellular biology* **7**, 379-387 (1987).

57 Rasheed, S., Nelson-Rees, W. A., Toth, E. M., Arnstein, P. & Gardner, M. B. Characterization of a newly derived human sarcoma cell line (HT-1080). *Cancer* **33**, 1027-1033 (1974).

58 Dominici, M. *et al.* Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy* **8**, 315-317, doi:10.1080/14653240600855905 (2006).

590 59 Mullers, E. *et al.* Novel Functions of Prototype Foamy Virus Gag Glycine- Arginine-
591 Rich Boxes in Reverse Transcription and Particle Morphogenesis. *J. Virol.* **85**, 1452-
592 1463, doi:10.1128/jvi.01731-10 (2011).

593 60 Trobridge, G., Vassilopoulos, G., Josephson, N. & Russell, D. W. Gene transfer with
594 foamy virus vectors. *Methods Enzymol* **346**, 628-648 (2002).

595 61 Barde, I., Salmon, P. & Trono, D. Production and titration of lentiviral vectors. *Current*
596 *protocols in neuroscience / editorial board, Jacqueline N. Crawley ... [et al.]* **Chapter 4**,
597 Unit 4.21, doi:10.1002/0471142301.ns0421s53 (2010).

598 62 Zeigler, M. *et al.* Prenatal diagnosis of Krabbe disease using a fluorescent derivative of
599 galactosylceramide. *Clinica chimica acta; international journal of clinical chemistry* **142**,
600 313-318 (1984).

601 63 Bligh, E. G. & Dyer, W. J. A rapid method of total lipid extraction and purification.
602 *Canadian journal of biochemistry and physiology* **37**, 911-917 (1959).

603 64 McAllister, R. G. *et al.* Lentivector integration sites in ependymal cells from a model of
604 metachromatic leukodystrophy: non-B DNA as a new factor influencing integration. *Mol*
605 *Ther Nucleic Acids* **3**, e187, doi:10.1038/mtna.2014.39 (2014).

606 65 Schambach, A. *et al.* Woodchuck hepatitis virus post-transcriptional regulatory element
607 deleted from X protein and promoter sequences enhances retroviral vector titer and
608 expression. *Gene Ther* **13**, 641-645 (2005).

609

610 Figure Legends

611 **Figure 1 – Schematic of FVVs used in this study.** All FVVs contain self-inactivating long-terminal
612 repeats (SIN-LTR, black boxes) with the U3 region deleted of promoter and enhancer activity. The cis-
613 acting sequences I and II (CASI/II) are viral sequences necessary for virion assembly. Promoter and
614 transgene of choice are inserted in a multiple cloning site (large white box). A post-regulatory element
615 (PRE) is included in all constructs to improve transgene expression. Insert abbreviations: PGK – murine
616 phosphoglycerate kinase promoter; GFP – enhanced green fluorescent protein; EFS – elongation factor 1
617 alpha short (intron-less version); ARSA – arylsulphatase A (codon optimised for human expression).

Figure 2 – Effect of FVV envelope on mMSC transduction efficiency and CPE. (a) The percent of GFP expressing mMSCs after 1 passage post-transduction with PGK-GFP enveloped with PFV Env (brick patterned) or SFV_{mac} Env (white) at different MOIs was determined by flow cytometry. The mean + SD of biological triplicates is shown. Significant differences between means at each MOI are indicated by asterisks according to the P value as determined by Two-Way ANOVA with Bonferroni's multiple comparisons test. (b-c) Photomicrographs of MSCs 20 hours after vector addition. Scale bar = 50 μm. (b) Transduced at MOI 30 using PFV Env; (c) Transduced at MOI 30 using SFV_{mac} Env.

Figure 3 – Comparison of the PGK and EFS promoter for sustained GFP expression in mMSCs. (a) The percentage of mMSCs expressing GFP and (b) their MFI following transduction at MOI 1 (squares) or MOI 30 (triangles) using FVV:PGK-GFP (white) or FVV:EFS-GFP (black). Flow cytometry analysis was performed when cells reached ~90% confluence at 1 to 10 passages post-transduction. For each passage, $1/10^{th}$ of the total cells were reseeded. Data points show the mean + SD of data from biological triplicates. Asterisks mark values that are significantly different from the previous passage of the same sample, as determined by Two-way ANOVA with Bonferroni multiple comparisons tests. (c-d) Representative photomicrographs of GFP fluorescence in mMSCs transduced at a MOI of 30 with PGK-GFP (c) or EFS-GFP (d), taken 3 passages post-transduction.

Figure 4 – FVV mediates high and sustained expression of ARSA in mMSCs. (a) Lysates of mu-MSC transduced cells were taken at each passage from 1 to 5 post-transduction. Lysates were mixed with σ-nitrocatechol sulphate in conditions that specifically enable hydrolysis to be catalysed by ARSA. The nmol of substrate hydrolysed per hour and per mg of total protein per ml are given as mean + SD of data from biological triplicates. Two-way ANOVA with Bonferroni multiple comparisons tests was used to identify significant differences between passages of the same transduced mMSCs. Where a passage was identified to be significantly different from the previous passage of the same population, asterisks are shown to indicate the P value. (b) The cell-culture medium of mMSCs transduced with the FVV indicated

was collected and used in the ARSA assay. The nmol of substrate hydrolysed per hour per ml of cell-culture medium is given as mean + SD of data from biological triplicates. (c) Human fibroblasts from 2 normal donors (N1 and N2) and 2 MLD patients' fibroblasts were loaded with BODIPY-sulphatide then cultured in the presence of dilute cell-culture medium from FVV:PGK-GFP (white) or FVV:PGK-ARSA (diamond patterned) transduced mMSCs. After 24 hours, the pmol of BODIPY-sulphatide present in the samples was determined by HPLC analysis. Values (mean + SD or technical triplicates) are normalised to BODIPY-lactosylceramide which was added immediately prior to lipid extraction as a recovery standard.

Figure 5 – mu-MSC characteristics are not perturbed by FVV transduction. (a) A panel of antibodies was used to test surface marker expression on mMSCs transduced or not with FVV. Representative plots (using mMSCs transduced with FVV:PGK-ARSA) are shown comparing isotype control antibody staining (blue) to staining with the antibody indicated above each plot (black). (b) The difference between isotype control and surface marker antibody signal intensity was quantified using the FlowJo Chi-squared T(x) comparison and plotted for untransduced MSCs (black), or mMSCs transduced with either FVV:PGK-GFP (white) or FVV:PGK-ARSA (diamond patterned). Higher values of T(x) show greater difference between surface marker antibody and the isotype control. (c-k) Representative photomicrographs are shown for mMSCs stained with Alizarin Red (c-e), Alcian Blue (f-h) or Oil Red O (i-k) without inducing differentiation (c,f,i) or after culturing in osteogenic (d-e), chondrogenic (g-h) or adipogenic (j-k) differentiation medium. mMSCs were untransduced or transduced with FVV-PGK-ARSA at a MOI of 30 as indicated.

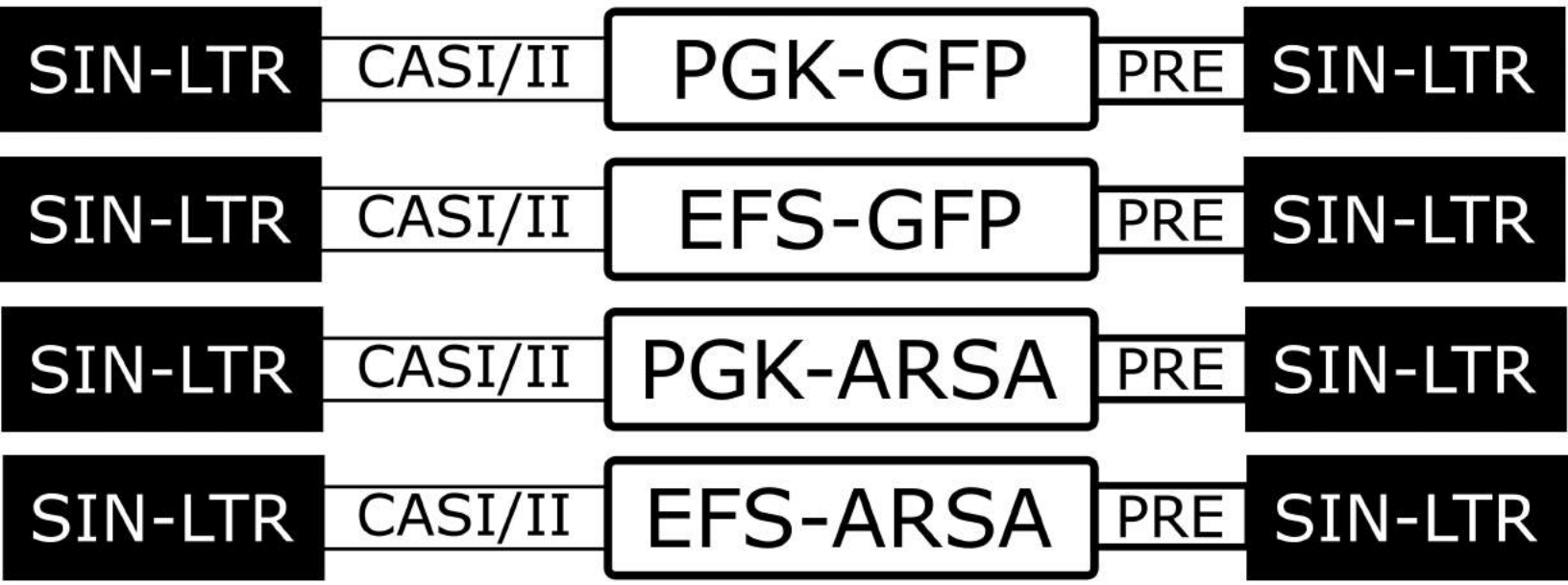
Figure 6 – FVV transduced mMSCs maintain their ability to graft in the brains of mice following intracerebroventricular delivery. (a) Sectioning of female mouse brains for qPCR analysis to determine long-term engraftment and distribution of male mMSCs 3 months after the direct delivery of 80 000 mMSCs into the right lateral ventricle (approximate location indicated with solid black circle). The dashed lines indicate the cuts made to produce 8 sections (numbered). (b) qPCR analysis to quantify

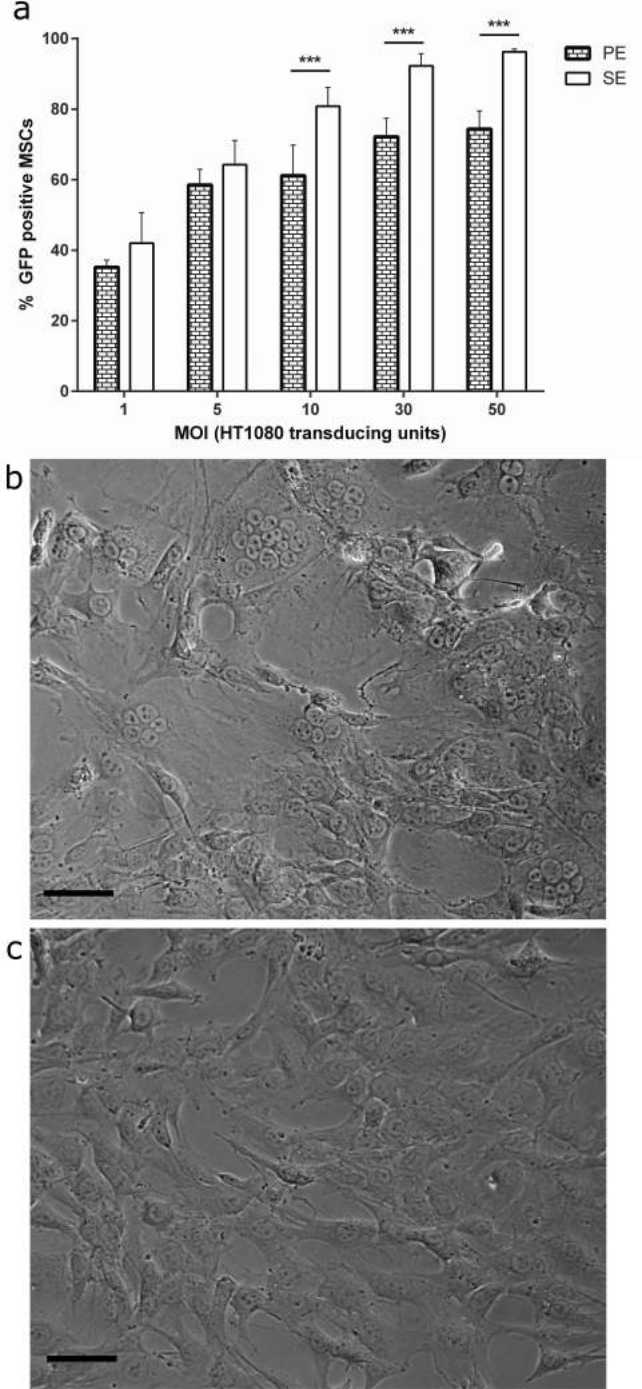
mMSCs in the brain sections depicted in a. Each data point represents a section from 1 mouse. White squares show the engraftment of untransduced mMSCs, whereas black diamonds show the engraftment of FVV transduced MSCs. Data for each group is from 6 treated mice. Only sections that were above the detection limit are shown. (c-f) Photomicrographs showing direct GFP fluorescence in mouse brains following injection of mMSCs transduced by FVV:PGK-GFP. Nuclei, stained with DAPI, are shown in blue. (c) Right lateral ventricle immediately post-injection; (d) Choroid plexus 45-days post-injection; (e) parenchyma showing needle track, 45-days post-injection; (f) glomerular layer of olfactory bulb, 45-days post-injection.

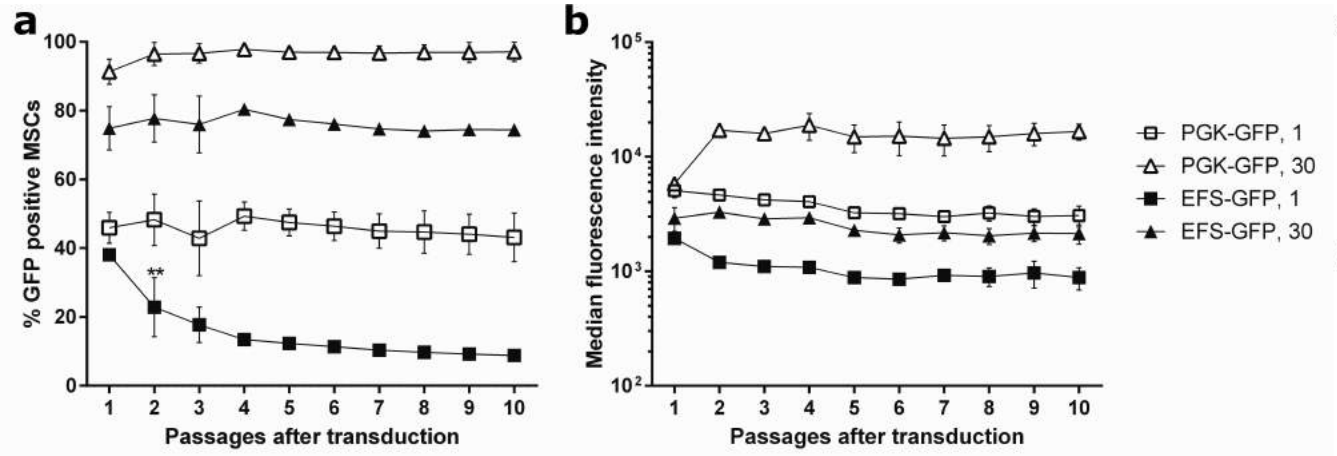
Figure 7 – Human MSCs can be efficiently and stably transduced by FVV. (a) hMSCs from 3 donors were transduced with FVV:PGK-GFP (white) or FVV:EFS-GFP (black) at the MOIs indicated on the x-axis and the percentage of GFP expressing cells was determined by flow cytometry 1 passage post-transduction. The same vectors, but using the PFV Env, were tested at MOI 50 (brick-patterned bars). (b-c) Photomicrographs of hMSCs 3 days post-transduction (no prior passages post-transduction) with FVV:PGK-GFP at a MOI of 100 showing normal cell morphology (b) and GFP fluorescence (c). (d-e) Each time >80% confluence was reached, $1/5^{th}$ of the hMSCs transduced at MOI 1 (squares) and MOI 100 (circles) were reseeded until 10 passages post-transduction. At each passage, the percentage of GFP expressing cells (f) and their median fluorescence intensity (g) was determined by flow cytometry. All panels show the mean + SD of data from biological triplicates. Asterisks in the x-axis labels indicates passages where cultures growth had slowed down, putatively entering senescence (7-10 post-transduction).

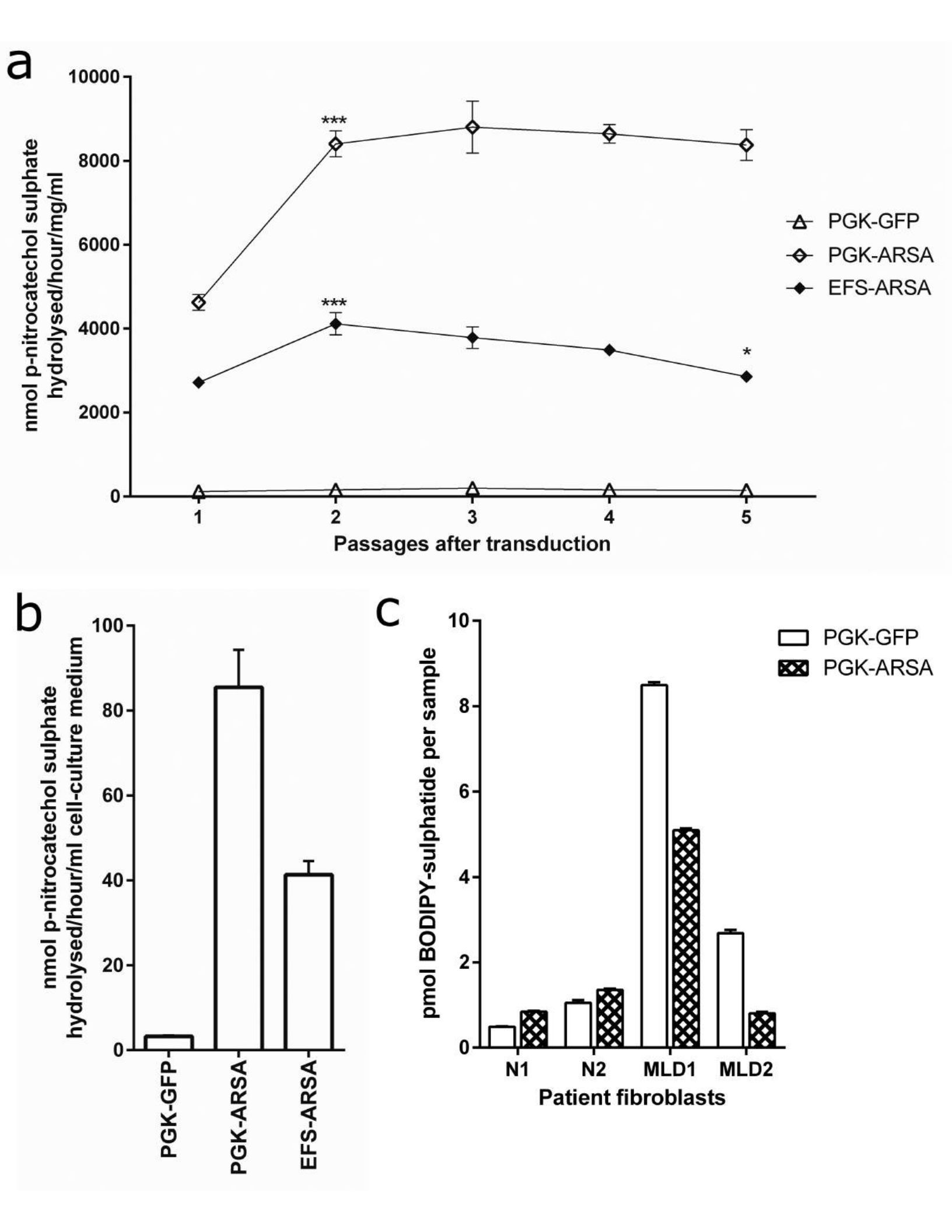
Figure S1 – Flow cytometry to determine percent of GFP expressing mMSCs. FlowJo plots of single cells in the 488-530/30 versus 488-610/20 channels are shown for untransduced mMSCs, FVV:EFS-GFP transduced at MOI 1 (showing the weakest GFP signal), and FVV:PGK-GFP transduced at MOI 30 (showing the strongest GFP signal) as indicated. The GFP negative and GFP positive gates were set

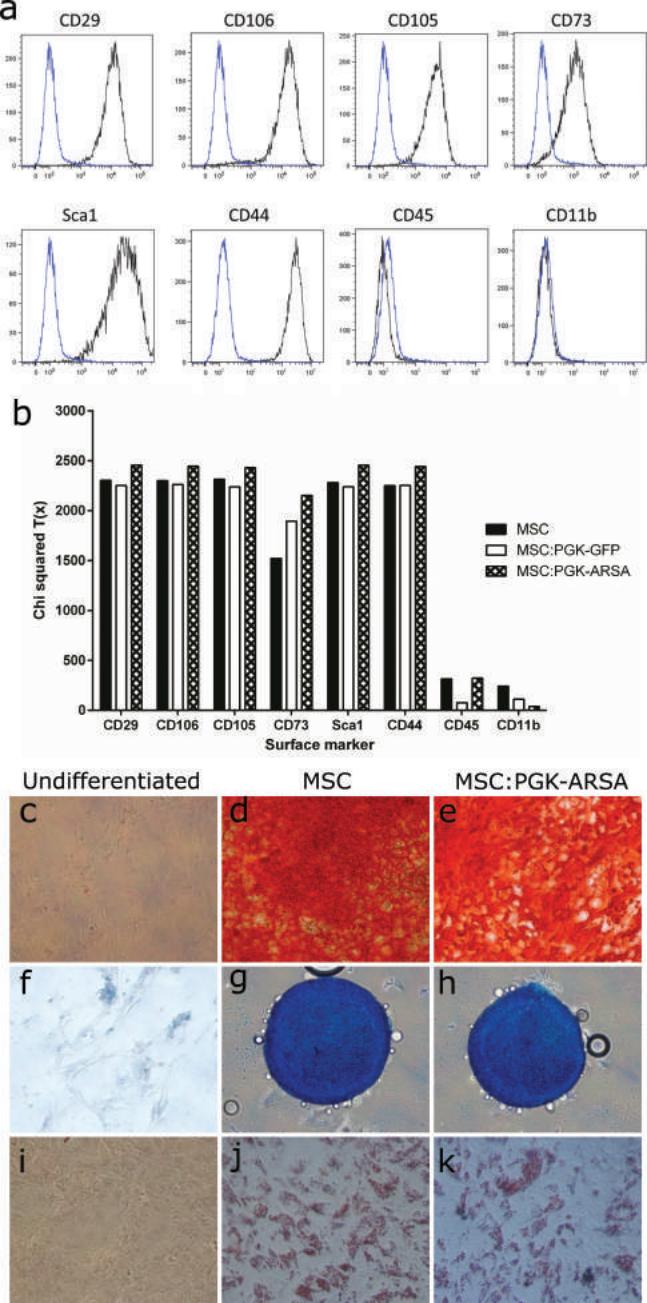
690 around the untransduced mMSC population and applied to transduced samples to determine the
691 percentage of GFP expressing mMSCs, as described in Materials and Methods.

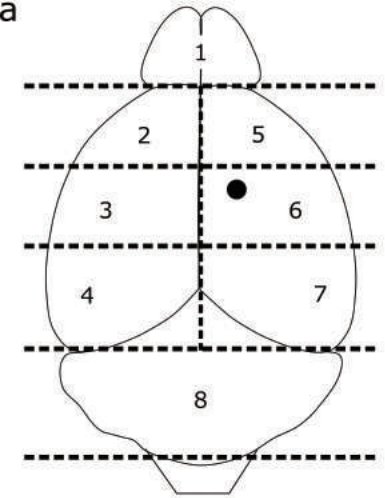

SIN-LTR
CASI/II
PGK-GFP
PRE
SIN-LTR
SIN-LTR
CASI/II
EFS-GFP
PRE
SIN-LTR
SIN-LTR
CASI/II
PGK-ARSA
PRE
SIN-LTR
SIN-LTR
CASI/II
EFS-ARSA
PRE
SIN-LTR

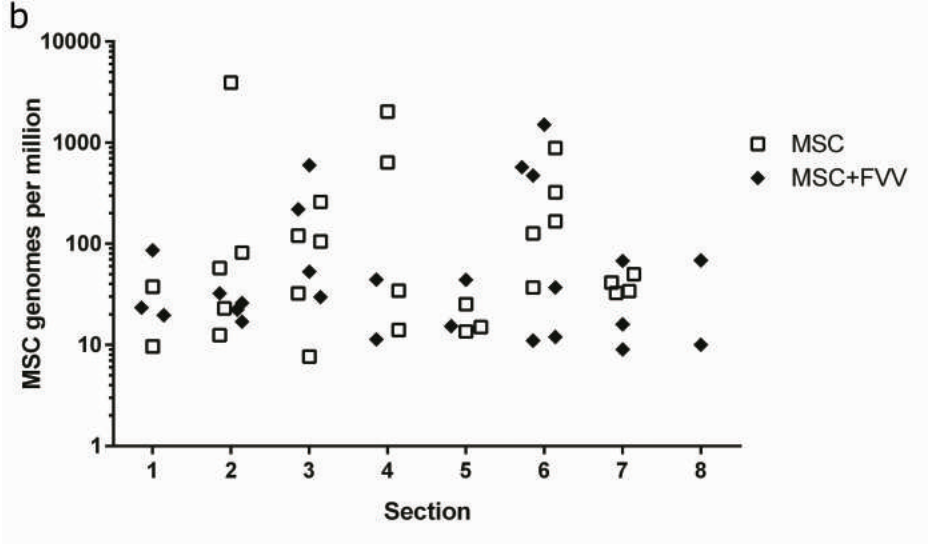

a
100
80
60
40
20
0
% GFP positive MSCs
***
***
***
PE
SE
1
5
10
30
50
MOI (HT1080 transducing units)
b
c

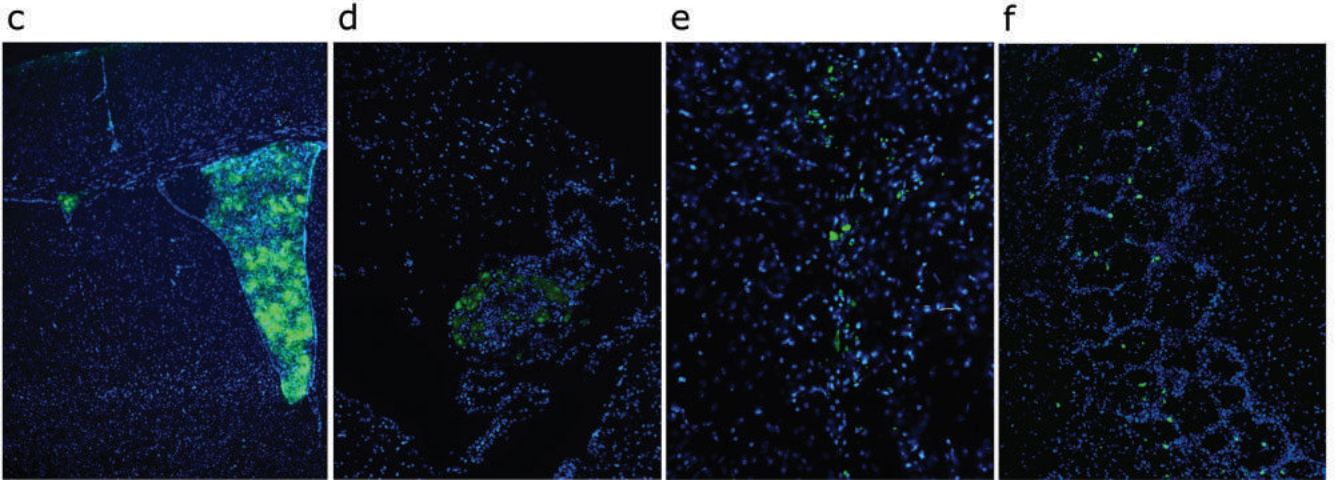

a
% GFP positive MSCs
100
80
60
40
20
0
**
1 2 3 4 5 6 7 8 9 10
Passages after transduction
b
Median fluorescence intensity
10^5
10^4
10^3
10^2
1 2 3 4 5 6 7 8 9 10
Passages after transduction
PGK-GFP, 1
PGK-GFP, 30
EFS-GFP, 1
EFS-GFP, 30
c
d

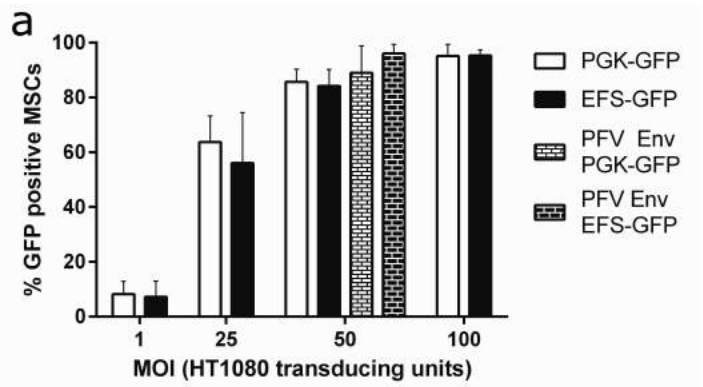

a
10000
8000
6000
4000
2000
0
nmol p-nitrocatechol sulphate hydrolysed/hour/mg/ml
***
***
*
1
2
3
4
5
Passages after transduction
PGK-GFP
PGK-ARSA
EFS-ARSA
b
100
80
60
40
20
0
nmol p-nitrocatechol sulphate hydrolysed/hour/ml cell-culture medium
PGK-GFP
PGK-ARSA
EFS-ARSA
c
10
8
6
4
2
0
pmol BODIPY-sulphatide per sample
PGK-GFP
PGK-ARSA
N1
N2
MLD1
MLD2
Patient fibroblasts

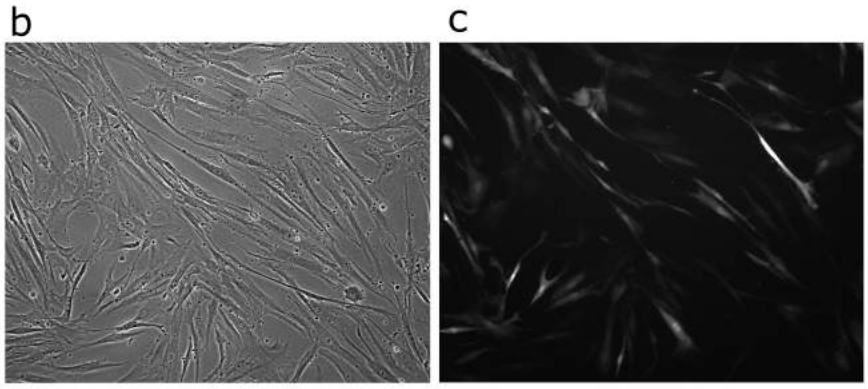

a
CD29
CD106
CD105
CD73
Sca1
CD44
CD45
CD11b
b
Chi squared T(x)
3000
2500
2000
1500
1000
500
0
MSC
MSC:PGK-GFP
MSC:PGK-ARSA
CD29
CD106
CD105
CD73
Sca1
CD44
CD45
CD11b
Surface marker
Undifferentiated
MSC
MSC:PGK-ARSA
c
d
e
f
g
h
i
j
k

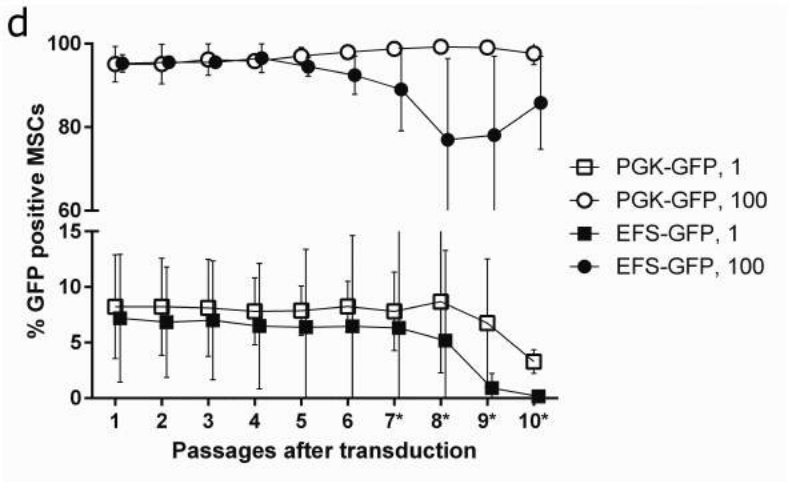

a
1
2
5
3
6
4
7
8

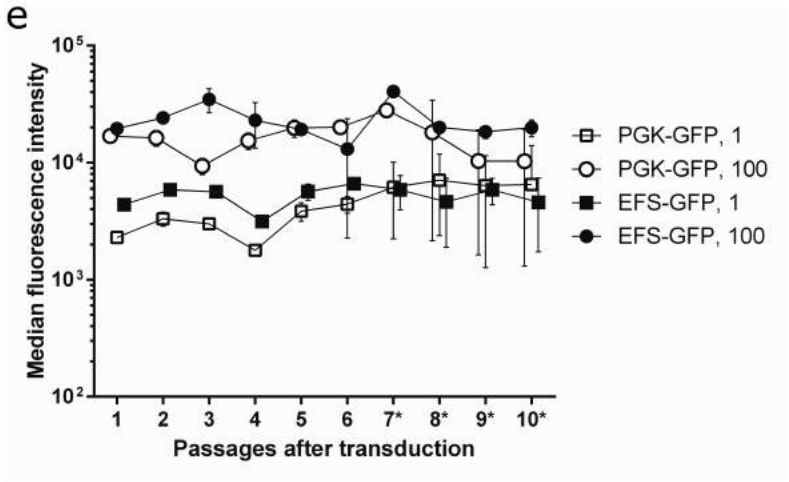

b
MSC genomes per million
10000
1000
100
10
1
1
2
3
4
5
6
7
8
Section
MSC
MSC+FVV


c
d
e
f


a
% GFP positive MSCs
100
80
60
40
20
0
1
25
50
100
MOI (HT1080 transducing units)
PGK-GFP
EFS-GFP
PFV Env PGK-GFP
PFV Env EFS-GFP


b
c


d
% GFP positive MSCs
100
80
60
15
10
5
0
1
2
3
4
5
6
7*
8*
9*
10*
Passages after transduction
PGK-GFP, 1
PGK-GFP, 100
EFS-GFP, 1
EFS-GFP, 100


e
Median fluorescence intensity
10^5
10^4
10^3
10^2
1
2
3
4
5
6
7*
8*
9*
10*
Passages after transduction
PGK-GFP, 1
PGK-GFP, 100
EFS-GFP, 1
EFS-GFP, 100