1   **A Safe Harbor-Targeted CRISPR/Cas9 Homology Independent Targeted Integration**

2   **(HITI) System for Multi-Modality Reporter Gene-Based Cell Tracking**

3

4   John J Kelly[1,2], Moe Saee-Marand[2], Nivin N Nyström[1,2], Yuanxin Chen[1], Melissa M Evans[1],

5   Amanda M Hamilton[1], John A Ronald[1,2*]

6

7   [1]Robarts Research Institute and [2]Dept. of Medical Biophysics, University of Western Ontario,

8   London, Ontario, N6A 5B7, Canada

9

10

11   Key words: CRISPR, reporter gene, HITI, cell tracking, AAVS1

12

13   Address correspondence to:

14   Dr. John A. Ronald

15   Robarts Research Institute,

16   Schulich School of Medicine and Dentistry,

17   University of Western Ontario,

18   London, Ontario, N6A 5B7,

19   Canada

20   Tel: 519-931-5777 x24391

21   Email: jronald@robarts.ca

22   **Abstract**

23   Imaging reporter genes can provide valuable, longitudinal information on the biodistribution,

24   growth and survival of engineered cells in preclinical models and patients. A translational

25   bottleneck to using reporter genes in patients is the necessity to engineer cells with randomly-

26   integrating vectors. CRISPR/Cas9 targeted knock-in of reporter genes at a genomic safe harbor

27   locus such as adeno-associated virus integration site 1 (AAVS1) may overcome these safety

28   concerns. Here, we built Homology Independent Targeted Integration (HITI) CRISPR/Cas9

29   minicircle donors for precise AAVS1-targeted simultaneous knock-in of fluorescence,

30   bioluminescence, and MRI (*Oatp1a1*) reporter genes. Our results showed greater knock-in

31   efficiency at the AAVS1 site using HITI vectors compared to homology-directed repair donor

32   vectors. Characterization of select HITI clones demonstrated functional fluorescence and

33   bioluminescence reporter activity as well as significantly increased Oatp1a1-mediated uptake of

34   the clinically-approved MRI agent gadolinium ethoxybenzyl diethylenetriamine pentaacetic acid.

35   As few as 10^6 Oatp1a1-expressing cells in a 50 µl subcutaneous injection could be detected *in*

36   *vivo* with contrast-enhanced MRI. Contrast-enhanced MRI also improved the conspicuity of both

37   sub-cutaneous and metastatic Oatp1a1-expressing tumours prior to them being palpable or even

38   readily visible on pre-contrast images. Our work demonstrates the first CRISPR/Cas9 HITI

39   system for knock-in of large DNA donor constructs at a safe harbor locus, enabling multi-modal

40   longitudinal *in vivo* imaging of cells. This work lays the foundation for safer, non-viral reporter

41   gene tracking of multiple cell types.

42

43   **Introduction**
44
45   Molecular-genetic imaging with reporter genes permits the *in vivo* visualization and tracking of

46   engineered cells, and thus, allows one to track the biodistribution, persistence, viability and in

47   some cases, activation state, of such cells[1,2]. Several reporter genes currently exist for visualizing

48   engineered cells using pre-clinical optical fluorescence imaging (FLI) and bioluminescence

49   imaging (BLI)[3–5] as well as those for clinical modalities such as magnetic resonance imaging

50   (MRI), positron emission tomography (PET) and photoacoustic imaging[6–8]. These non-invasive

51   cell tracking tools are invaluable for understanding mechanisms of disease progression and the

52   evaluation of treatments in pre-clinical animal models. Important examples in cancer research

53   include the tracking of therapeutic stem cells[9–11], tracking immune cell migration, cancer

54   progression and metastasis[12–14], and evaluating tumour response to novel anticancer

55   therapeutics[15,16]. More recently, the use of reporter genes to track therapeutic cells has been

56   translated into the clinic. In this case, cytotoxic T cells were engineered to express a chimeric

57   antigen receptor to target glioma cells, as well as a herpes simplex virus type 1 thymidine kinase

58   (HSV1-TK) dual reporter-suicide gene (that selectively uptakes the PET tracer [^{18}F]FHBG) to

59   track the localization and viability of the injected therapeutic cells in glioma patients[17,18].

60

61   Although reporter genes have great potential for therapeutic cell tracking, their functionality is

62   best utilized when the genes are stably integrated into the desired cell's genome – allowing

63   reporter gene expression throughout the lifetime of the cell and in any subsequent daughter cells.

64   Retroviral vectors, such as those derived from HIV lentiviruses, have generally been used for

65   transgene integration due to their high transfection efficiency, large transgene capacity and their

66   ability to transduce a variety of dividing and non-dividing cell types. However, the slow

67   acceptance of using reporter genes for tracking cell-based therapies may in part be due to the

68   increased risk of random or quasi-random insertional mutagenesis when transgenes are delivered

69   using viral vectors[19]. Indeed, in previous clinical trials involving children with X-linked severe

70   combined immunodeficiency, a Moloney murine leukemia virus–based γ-retrovirus vector

71   expressing the interleukin-2 receptor γ-chain (γc) complementary DNA successfully restored

72   immunity in most patients. However, 5 of the 20 patients also developed leukemia, of which one

73   child died, as a result of insertional mutagenesis and transactivation of proto-oncogenes[20–22]. For

74   future cell-based therapies, it is therefore highly desirable to edit cells with reporter genes in a

75 safe and site-specific manner. The application of such editing tools would allow for longitudinal

76 cell tracking to confirm that the cells are performing their intended role, as well as to detect any

77 ectopic growths or misplaced targeting at the earliest time point. This will ultimately give the

78 clinician greater control and confidence in the outcomes of the targeted therapy.

79

80 Genomic safe harbors can incorporate exogenous pieces of DNA and permit their predictable

81 function, but do not cause alterations to the host genome, nor pose a risk to the host cell or

82 organism[23]. Several studies have successfully used genome editing tools such as zinc finger

83 nucleases (ZFN) and transcription activator-like effector nucleases (TALENs) to incorporate

84 reporter genes at the adeno-associated virus integration site 1 (AAVS1) safe harbor locus, with

85 no detrimental effects[24–26]. Although ZFN and TALENs have shown great promise as targeted

86 DNA editors, they are time consuming, expensive and challenging to engineer as new nuclease

87 sequences must be generated for every new genomic target. Alternatively, clustered regularly

88 interspaced short palindromic repeat/Cas9 (CRISPR/Cas9), which was developed by several

89 groups in 2013 [27–30], allows for quicker, cheaper and easier-to-design human genome editing.

90 CRISPR/Cas9 utilizes short guide RNAs (gRNAs; ~20bp in length) to direct the Cas9

91 endonuclease to a specific genomic locus and induce a double stranded DNA break. Both the

92 Cas9 enzyme and gRNA sequences can be encoded in a single plasmid and when co-transfected

93 with a donor DNA plasmid can lead to higher homology directed repair (HDR) knock-in

94 efficiency than previous editing tools[31]. We have previously described the first CRISPR-Cas9

95 system for AAVS1 integration of donor constructs containing an antibiotic resistance selection

96 gene and both fluorescence (*tdTomato*) and bioluminescence (*Firefly luciferase*) reporter

97 genes[32]. We were able to confirm the correct and stable integration of donor DNA at the AAVS1

98 site and functional reporter gene expression *in vivo*. However, some of the limitations of our

99 study include (i) the low editing efficiency (~3.8%) of HEK-293T cells; (ii) the use of large

100 CRISPR/Cas9 and donor DNA plasmids that contained bacterial and antibiotic resistance genes –

101 which limit transfection efficiency and would have associated safety concerns for clinical

102 translation; and (iii) the lack of a translationally relevant reporter gene. In this study, we aimed to

103 address these limitations by improving the efficiency and clinical safety of reporter gene

104 integration at the AAVS1 safe harbor site and included a translationally relevant reporter gene.

105

106   We posited that the low editing efficiency of our first system was due in part to reduced

107   transfection and knock-in efficiency, which is common with larger DNA plasmids, and the use of

108   the HDR repair pathway for integration, which is intrinsically inefficient and not readily

109   accessible to non-dividing cells[33]. In contrast to HDR-mediated DNA repair, the non-

110   homologous end joining (NHEJ) pathway is active in both proliferating and non-proliferating

111   cells and is generally considered more efficient than HDR in mammalian cells[34]. Recent studies

112   have shown that by designing a CRISPR/Cas9 system that includes the same gRNA cut site in

113   the donor vector as the genomic target site, the NHEJ repair pathway will more efficiently lead

114   to transgene integration in zebrafish[35] and mammalian cells[36,37]. Suzuki *et al.* refer to this

115   mechanism as homology independent targeted integration (HITI), which is expected to lead to

116   increased insertion in the forward rather than the reverse direction, as intact gRNA target

117   sequences will be preserved in the latter[38]. Therefore, we postulated that HITI will increase the

118   efficiency of reporter gene integration at the AAVS1 site (Figure 1A), compared to HDR. To

119   address the problem of size and bacterial/antibiotic resistance genes in plasmids, our group and

120   Suzuki *et al.*[38], previously designed minicircles (MC) to express genes of interest[39,40]. First

121   described by Darquet *et al.*[41], MCs lack the bacterial backbone and antibiotic resistance genes

122   that would otherwise compromise biosafety and clinical translation. In addition, the removal of

123   the prokaryotic backbone also greatly reduces the size of the vector, thus improving transfection

124   efficiency or providing space for the inclusion of other transgenes. To that end, we aimed to

125   improve on our previous work by including a translationally relevant reporter gene in a multi-

126   modality imaging HITI MC donor. We determined that the rat organic-anion-transporting

127   polypeptide 1A1 (*Oatp1a1*) gene was an ideal candidate. *Oatp1a1* is a positive contrast magnetic

128   resonance imaging (MRI) reporter gene due to its ability to uptake a clinically approved, liver-

129   specific paramagnetic contrast agent called gadolinium ethoxybenzyl diethylenetriamine

130   pentaacetic acid (Gd-EOB-DTPA; Primovist/Eovist)[42]. We have previously shown that *Oatp1a1*

131   is a sensitive, quantitative, MRI reporter for 3-dimensional cancer cell distribution *in vivo*[43]. The

132   purpose of this study was to develop HITI MC donor vectors for CRISPR/Cas9 editing of cells at

133   the AAVS1 locus with three reporter genes to allow for multimodality, longitudinal *in vivo*

134   monitoring of their fate following transplantation.

135

136   **Materials and Methods**

137

138 Constructs

139 Construct designs are shown in Figure 1B. The pCas9-AAVS1guideRNA-zsG-MC (Cas9-

140 AAVS1-MC) and pCas9-scrambledRNA-zsG-MC (Cas9-scrambled-MC) parental plasmids

141 originated from pCas-Guide-AAVS1 and pCas-Guide-Scrambled plasmids purchased from

142 Origene (Maryland, USA). The Cas9 enzyme and guide RNA sequences were cloned between

143 *attB* and *attP* recombination sites in a minicircle bacterial backbone containing a ZsGreen (*zsG*)

144 fluorescence reporter driven by the elongation factor 1-α promoter (*hEF1α*). The AAVS1-HDR-

145 tdT-Fluc2-Oatp1a1-MC (HDR-MC) construct was derived from an HDR vector lacking the

146 *Oatp1a1* gene as we described previously[32]. This plasmid is driven by the *hEF1α* promoter,

147 expresses tdTomato (*tdT*), firefly luciferase (*Fluc2*) and Organic anion transporting polypeptide

148 1a1 (*Oatp1a1*) using a self-cleaving 2a peptide system. For improved expression, the plasmids

149 also contain the woodchuck hepatitis virus post-transcriptional regulatory element (*WPRE*)

150 followed by the human growth hormone polyadenylation signal (*hGH* polyA). The HDR plasmid

151 contains the left and right homologous arms (RHA:527bp, LHA:481bp) that are complementary

152 to the region flanking the AAVS1 cut site; the homologous arms were obtained from the

153 pAAVS1-puroDNR plasmid from Origene (Maryland, USA). The *Oatp1a1* gene was added

154 through PCR amplification from a previously made vector we constructed using

155 PGK_Straw_E2A_Oatp1a1 (a kind gift from Dr. Kevin Brindle's laboratory; University of

156 Cambridge). Using the HDR-MC parental plasmid as the template, we generated the pAAVS1-

157 HITI-tdT-Fluc2-Oatp1a1-MC (HITI-MC) parental plasmid using the In-Fusion cloning kit from

158 Clontech (Takara Bio, California, USA). Using restriction enzyme digestion, we extracted the

159 bacterial backbone and minicircle recombination sites and then extracted the three reporter genes

160 (without the homologous arms); *tdT*, *Fluc2* and *Oatp1a1* from the HDR-MC construct using

161 PCR. However, for HITI functionality we designed our primers to also include a 23 bp extension

162 (5'-GTTAATGTGGCTCTGGTTCTGGG-3') downstream of the polyA sequence, which

163 incorporates the same cut site and protospacer adjacent motif (PAM) sequence for our AAVS1

164 gRNA which allows for Cas9 cutting of both the MC and genomic DNA.

165

166 Minicircle Production

167 ZYCY10P3S2T E. coli (System Biosciences, Palo Alto, CA, USA) were transformed with the

168 original parental plasmids (PP) of all four constructs; HDR-MC or HITI-MC, Cas9-AAVS1-MC

169   and Cas9-scrambled-MC and viable colonies selected using kanamycin plates. Colonies were

170   picked 24 hrs after transformation, grown in 6 ml of lysogeny broth (LB) with kanamycin for 6

171   hrs at 37ºC, followed by growth in terrific broth (TB) for 12 hrs at 37ºC. To induce expression of

172   the phiC31 integrase for MC production via *attB* and *attP* recombination, 100 ml of LB broth

173   together with 100 µl of 20% arabinose induction solution (System Biosciences, Palo Alto, CA,

174   USA) and 4 ml of 1N NaOH was added to the culture and grown for 5.5 hrs at 30°C. An

175   endotoxin-free maxi kit (Qiagen, Valencia, CA, USA) was used to purify both PP and MC.

176   Following purification of the MC products, PP contamination was removed using the Plasmid

177   Safe ATP-dependent DNase kit (Epicentre, WI, USA), and the products where cleaned and

178   concentrated using the Clean & Concentrator-25 kit (Zymo Research, CA, USA).

179

180   Cell Culture and Transfection

181   Human embryonic kidney 293T cells and human adenocarcinoma HeLa cells (both from ATCC,

182   Manassas, VA, USA) were grown in DMEM medium (Wisent Bioproducts, Québec, Canada)

183   supplemented with 10% Fetal Bovine Serum (FBS; Wisent Bioproducts, Québec, Canada) and

184   1x Antibiotic-Antimycotic (ThermoFisher, Waltham, MA, USA). Human grade 4

185   adenocarcinoma PC3 cells were a kind gift from Dr. Hon Leong (Western University, ON,

186   Canada) and were grown in RPMI (Wisent Bioproducts, Québec, Canada) supplemented with

187   5% FBS and 1x Antibiotic-Antimycotic. Cells were transfected with the linear polyethylenimine

188   transfection agent jetPEI (Polyplus-transfection, Illkirch, France), according to the

189   manufacturer's instructions. Briefly, cells were grown in 6-well plates until 80-90% confluency

190   and co-transfected with 1 µg each of Cas9-AAVS1-MC or Cas9-Scrambled-MC together with 1

191   µg of the donor MC constructs: HDR-MC or HITI-MC, for a total DNA mass of 2 µg. The DNA

192   was prepared in 150 mM NaCl and complexed with 4 µl of jetPEI reagent per well.

193

194   FACS and Flow Analysis

195   All FACS and flow cytometry was performed at the London Regional Flow Cytometry Facility

196   (Robarts Research Institute, London, Canada). Forty-eight hours post transfection, the population

197   of cells displaying both red (tdTomato) and green (zsGreen) fluorescence were sorted using a BD

198   FACSAria III cell sorter (BD Biosciences, San Jose, CA, USA). At selected timepoints

199   following FACS the cells were analyzed for tdTomato fluorescence using a FACSCanto flow

200   cytometer (BD Biosciences, San Jose, CA, USA). Either 14- or 21-days post the initial sort, the

201   cells were again sorted on the FACSAria III to purify tdTomato positive cells only (referred to as

202   the pooled population). In this regard, our protocol aimed to sort cells that had incorporated the

203   MC inserts (based on tdTomato fluorescence) into the genome and excluded any cells that had

204   randomly integrated Cas9 MC DNA (zsGreen). At the same time as the second (tdTomato) sort,

205   individual cells were plated into wells of a 96-well plate to enable single cell colonies to be

206   grown and expanded (referred to as clonal cell populations).

207

208   <u>Genomic DNA Extractions and AAVS1 Integration Analysis</u>

209   Extraction of genomic DNA from the pooled population of cells was performed using the

210   DNeasy Blood and Tissue kit (Qiagen, Valencia, CA, USA) following manufacturer's

211   instructions. DNA quality and concentrations were measured on a NanoDrop 1000

212   spectrophotometer (ThermoFisher). Extraction of genomic DNA from clonal populations was

213   performed as we described previously[32]. Briefly, cell pellets were resuspended in a QuickExtract

214   DNA extraction solution (Lucigen, Middleton, WI, USA), incubated at 65°C for 10 mins,

215   vortexed and incubated at 98°C for 5 mins. The DNA was then directly used for PCR or stored at

216   -20°C. To check for integration at the AAVS1 site, two primers where designed to amplify the 3'

217   junction between the donor cassette and the AAVS1 site outside of the homologous arm region.

218   The forward primer was uniquely complementary to the polyA tail in the MC cassette (5'-

219   CCTGGAAGTTGCCACTCCAG-3') and the reverse primer to the AAVS1 site (5'-

220   AAGGCAGCCTGGTAGACAGG-3'). A 1.3 kb PCR product was produced if the MC-HDR

221   was correctly integrated at the AAVS1 site and a 1.7 kb PCR product if MC-HITI was correctly

222   integrated. GAPDH primers were designed as DNA loading controls and to confirm successful

223   DNA extractions: forward 5'-TTGCCCTCAACGACCACTTT-3' and reverse 5'-

224   GTCCCTCCCCAGCAAGAATG-3' and yielded a PCR product of 502 bp. Agarose gel

225   electrophoresis with 1% agarose gels and RedSafe (FroggaBio, ON, Canada) was used to

226   separate and visualize PCR products.

227

228   <u>*In Vitro* Fluorescence and Bioluminescence Imaging</u>

229   The pooled and clonal cell populations were evaluated for tdTomato fluorescence expression on

230   an EVOS FL auto 2 microscope (ThermoFisher, Waltham, MA, USA). For BLI experiments,

231 varying cell numbers were plated in triplicates into black walled 96-well plates. D-luciferin (0.1

232 mg/ml; PerkinElmer, Waltham, MA, USA) was added to each well and images rapidly collected

233 on an IVIS Lumina XRMS In Vivo Imaging System (PerkinElmer) equipped with a cooled CCD

234 camera. Average radiance values in photons/sec/cm^2/steradian were measured from regions of

235 interest drawn around each well using LivingImage software (PerkinElmer).

236

237 *In Vitro* Magnetic Resonance Imaging

238 Naïve and Oatp1a1-expressing cell clones (2×10^6) were seeded in T-175 flasks and grown for 3

239 days. Cells were incubated with media containing 6.4 mM Gd-EOB-DTPA or with media

240 containing an equivalent volume of PBS for 90 minutes at 37°C and 5% CO_2. Cells were then

241 washed 3 times with PBS, trypsinized and pelleted in 0.2 mL Eppendorf tubes, and placed into a

242 2% agarose phantom mould that was incubated in a 37°C chamber for two hours to mimic body

243 temperature. MRI was performed on a 3-Tesla GE clinical MRI scanner (General Electric

244 Healthcare Discovery MR750 3.0 T, Milwaukee, WI, USA) and a 3.5-cm diameter birdcage RF

245 coil (Morris Instruments, Ottawa, ON, Canada). A fast spin echo inversion recovery (FSE-IR)

246 pulse sequence was used with the following parameters: field of view (FOV) = 256 x 256,

247 repetition time (TR) = 5000 msec, echo time (TE) = 19.1 msec, echo train length (ETL) = 4,

248 number of excitations (NEX) = 1, receiver bandwidth (rBW) = 12.50 MHz, inversion times (TI)

249 = 50, 100, 125, 150, 175, 200, 250, 350, 500, 750, 1000, 1500, 2000, 2500, 3000 msec, in-plane

250 resolution = 0.27 mm^2, slice thickness = 2.0 mm, scan time = 5 min and 25 sec per inversion

251 time. Spin-lattice relaxation rates were computed via MatLab by overlaying the image series and

252 calculating the signal intensity on a pixel-by-pixel basis across the inversion time image series,

253 followed by a fitting of the data into the following equation to output the spin-lattice relaxation

254 time, where S represents signal intensity, κ represents the scaling factor, and ρ represents proton

255 spin density:

256 $S = \kappa \cdot \rho\,(1 - 2 \cdot e\,^\wedge\,((-TI)/T1) + e\,^\wedge\,((-TR)/T1))$

257

258 Animal models

259 All animal protocols were approved by the University Council on Animal Care at the University

260 of Western Ontario (Protocol #2015-058) and follow the Canadian Council on Animal Care

261 (CCAC) and Ontario Ministry of Agricultural, Food and Rural Affairs (OMAFRA)

262   guidelines. Crl:NU-*Foxn1*[nu] (nude) male mice (Charles River Laboratories, Wilmington, MA,

263   USA; N = 3-5) aged 6-8 weeks were used for subcutaneous tumour model injections and

264   NOD.Cg-*Prkdc*[scid] *Il2rg*[tm1WjI]/SzJ (NSG) immunodeficient male mice (obtained from the

265   Humanized Mouse and Xenotransplantation Facility at the Robarts Research Institute, University

266   of Western Ontario, London, Canada; N = 3) for experimental metastasis models (intravenous

267   cell injections).

268

269   *In Vivo* Bioluminescence Imaging

270   BLI was performed on the same IVIS Lumina XRMS system described for *in vitro* imaging.

271   Mice were anesthetized with 2% isoflurane in 100% oxygen using a nose cone attached to an

272   activated carbon charcoal filter for passive scavenging and kept warm on a heated stage. BLI

273   images were acquired with automatic exposure times until the peak BLI signal was obtained (up

274   to 40 mins). Regions of interest were manually drawn using LivingImage software to measure

275   average radiance (photons/sec/cm^2/sr). The peak average radiance was used for quantification for

276   each mouse.

277

278   *In Vivo* Magnetic Resonance Imaging and Quantification

279   All mouse MRI scans were performed with a custom-built gradient insert and a bespoke 5 cm

280   diameter solenoidal RF coil, as we described previously[43]. Mice were kept anaesthetized during

281   the scan with 2% isoflurane administered via a nose cone attached to the coil. T_1-weighted

282   images were acquired using a spoiled gradient recalled acquisition in steady-state pulse sequence

283   using the following parameters: field of view, 50 mm; repetition time, 14.7 msec; echo time, 10.5

284   msec; receiver bandwidth, 31.25 MHz; echo train length, 4; frequency and phase, 250 x 250; flip

285   angle, 60 degrees; number of excitations, 3; 200 μm isotropic voxels; scan time, approximately

286   15 minutes per mouse. Pre-contrast images were acquired followed by administration of 1.67

287   mmol/kg Gd-EOB-DTPA (Primovist; Bayer, Mississauga, ON, Canada) via the tail vein. Mice

288   were then re-imaged 20 minutes later for immediate post-contrast images, which provide positive

289   contrast to many tissues, including the naïve tumours, as a result of Gd-EOB-DTPA pooling;

290   and/or 5 hours later for Oatp1a1-specific uptake. This time-point was determined to allow

291   enough time for Gd-EOB-DTPA to be cleared, yet still provide strong positive contrast in

292   Oatp1a1-expressing cells[42,43]. Contrast-to-noise ratio (CNR) measurements were calculated from

293 MR images using ITK-snap open source software (www.itksnap.org)[44]. Tumours were manually

294 segmented in 3 dimensions by tracing the tumour or control tissue (hind leg muscle) with

295 polygon and paintbrush tools and pixel intensity recorded in every slice. The CNR of tumours

296 was calculated by taking the signal intensity of the difference between tumour regions and

297 muscle tissue divided by the standard deviation of background signal

298 $(CNR = \frac{attenuation_{tumour} - attenuation_{muscle}}{Std.Dev._{background}})$.

299

300 Oatp1a1-induced Gd-EOB-DTPA uptake MRI and BLI sensitivity

301 To evaluate the cellular detection sensitivity of Oatp1a1 expressing cells with Gd-EOB-DTPA-

302 enhanced MRI, nude mice were injected with 50 µl of cell suspensions in PBS containing $3x10^6$

303 total cells/injection at the following ratios: $3x10^6$ naïve cells alone; 10^4 PC3-HITI + $2.99x10^6$

304 naïve cells; 10^5 PC3-HITI + $2.9x10^6$ naïve cells; 10^6 PC3-HITI + $2x10^6$ naïve cells; and $3x10^6$

305 PC3-HITI cells alone, subcutaneously in 5 locations on the back/flank region. Immediately after

306 cell injections, 1.67 mmol/kg Gd-EOB-DTPA was injected into the tail vein and mice were

307 imaged on a 3T clinical grade MR scanner 5 hours later. This time point allows for clearance of

308 Gd-EOB-DTPA from the body yet provides sufficient time for the agent to penetrate the

309 subcutaneous injections sites and accumulate in Oatp1a1 expressing cells. After MRI, mice were

310 moved to the IVIS scanner and injected with 100 µl of 30 mg/ml D-Luciferin intraperitoneally

311 and BLI was performed, as described earlier.

312

313 293T and PC3 tumour models

314 293T or PC3 naïve and HITI engineered cells were injected subcutaneously ($2.5x10^6$ 293Ts and

315 $1x10^6$ PC3s) on the left and right flanks of nude mice, respectively (293T, N = 2; PC3, N = 5).

316 For experimental metastasis studies, $5x10^5$ PC3 naïve or HITI engineered cells were injected into

317 the tail veins of NSG mice (N = 3). Tumour growth was tracked on a weekly basis with BLI, as

318 described above. MRI was performed on mice at various timepoints, as indicated in the results

319 section. Firstly, a pre-contrast scan was performed on all mice, followed immediately with

320 injection of the Gd-EOB-DTPA contrast agent into the tail vein (1.67 mmol/kg). For some

321 experiments the mice were re-scanned 15-20 mins after contrast injection to show tumour and

322 whole-body distribution of Gd-EOB-DTPA. In all instances MRI scans were performed ~5 hours

323 post-contrast injection since Oatp1a1 expressing cells still retain Gd-EOB-DTPA and show

324   strong positive contrast at this time point. This also allows enough time for washout of Gd-EOB-

325   DTPA in most tissues and organs (except for the gastrointestinal tract and bladder where cleared

326   Gd-EOB-DTPA accumulates before being excreted)[42].

327

328   <u>Statistics</u>

329   Statistical analysis was performed with GraphPad Prism version 7 (GraphPad Software Inc., CA,

330   USA; www.graphpad.com) software. One-way ANOVA with Tukey's multiple comparison test

331   was used for *in vitro* and *in vivo* BLI and CNR data analysis. An unpaired one-tailed *t*-test was

332   used to analyze the increase in CNR for PC3-HITI day 11 vs day 46 tumours.

333

334   **Results**

335   <u>CRISPR/Cas9 engineering of multiple human cell types with tri-modal reporter gene minicircle</u>

336   <u>donors</u>

337   In this study, we designed our trimodal reporter gene system in MC constructs to reduce the size

338   and immunogenicity of our donor DNA and to remove antibiotic resistance genes. To compare

339   the efficiency of HDR vs HITI editing at the AAVS1 site, we designed two donor and two Cas9-

340   expressing MCs, as shown in Figure 1B. The HDR and HITI constructs were engineered to

341   express tdTomato (*tdT*), firefly luciferase (*Fluc2*) and rat organic anion transporting polypeptide

342   1a1 (*Oatp1a1*) genes under the control of an *EF1α* promoter and 2A self cleaving peptide system

343   (Figure 1B). The HDR and HITI parental plasmids initially measured 11.9 kb and 10.4 kb in

344   size, which was then reduced to 7.9 kb and 6.4 kb when recombined into MCs, respectively, as

345   confirmed by agarose gel electrophoresis (Figure 1C). The HDR-MC was flanked by left and

346   right AAVS1 homologous arms either side of the AAVS1 genomic cut site, whereas the HITI

347   donor contained the same CRISPR/Cas9 cut site as the AAVS1 genomic site (Supp. Figure 1). In

348   this instance, if the MC DNA inserted into the correct orientation at the AAVS1 site, the

349   CRISPR/Cas9 cut sites would be lost and the trimodal reporter genes would be stably integrated

350   into the genome (Supp. Figure 1). The Cas9 expressing MCs were designed to contain the

351   necessary RNA scaffolding and gRNA sequences targeting the AAVS1 site or a scrambled

352   gRNA control, alongside a zsGreen (*zsG*) fluorescent reporter gene (Figure 1B). Both the pCas9-

353   AAVS1-MC and pCas9-scrambled-MC constructs measured 12.5kb in parental plasmid form,

354   and 8.6kb in MC form (Figure 1B).

355

356   Our first objective was to determine the correct integration of our donor MCs in three human cell

357   lines; HEK 293T, HeLa and PC3 cells. All 3 cell lines where co-transfected with the HDR-MC

358   or HITI-MC together with either the Cas9-AAVS1-MC or Cas9-scrambled-MC (as outlined in

359   Figure 1D) and grown for 48hrs. The cells were then FACs sorted for tdT+/zsG+ cells in order to

360   purify cells that were successfully co-transfected, and tdT fluorescence was then tracked every 7

361   days using flow cytometry (Supp. Figure 2A-B). In two separate experimental groups, the cells

362   were then resorted 14 or 21 days later for tdT+/zsG- cells, to ensure that the cell populations had

363   not randomly integrated the Cas9-zsG MCs into the genome (Supp. Figure 2C). Both PC3

364   experimental groups were resorted 14 days after the initial sort (and not 21 days later) due to

365   lower transfection rates. However, resorting the cells 14 or 21 days later had a negligible effect

366   on tdT+ cell populations across the timepoints. For almost all cell types there was a higher

367   percentage of tdT+ fluorescence cells at endpoint in the HITI-AAVS1 groups (pink shading,

368   Supp. Figure 2C), suggesting better or more stable integration compared to HDR-AAVS1

369   groups.

370

371   <u>Mixed cell population (MCP) integration and BLI analysis</u>

372   We next performed junctional PCR analysis on extracted DNA samples to determine whether the

373   tdT+ MCPs had correctly incorporated the large, trimodal donor MCs into the AAVS1 site in the

374   right orientation (Supp. Figure 3A). A correct integration band (1.4 kb) was detected for all

375   HITI-guideAAVS1 (HITIgA) engineered cells (very low transfection efficiency for PC3 cells

376   may explain why the integration band was weak) as well as a correct integration band (1.3 kb)

377   for HDR-guideAAVS1 (HDRgA) cells for 293T and HeLa MCPs. There were no integration

378   bands for the control naïve cells or cells engineered with scramble guide RNA (HITI/HDRgS).

379   Next, we performed *in vitro* BLI experiments to determine if the integrated reporter gene was

380   functioning in the MCPs. Varying numbers of each cell type were imaged with BLI after

381   addition of D-luciferin to visualize *FLuc2* expression (Suppl Figure 3B). In all cell types, there

382   was a positive correlation between BLI signal and cell number and, interestingly, a consistently

383   higher signal was seen in all cell types engineered with HITI-guideAAVS1 compared with HDR-

384   guideAAVS1.

385

386   <u>HITI is more efficient than HDR in clonal populations</u>

387   Next, we used clonal cell isolation to determine whether HITI or HDR was more efficient at

388   correctly integrating our large donor MCs at the AVVS1 site. Single cell tdT+ clones were

389   isolated from the 293T and PC3 MCPs into 96-well plates during a third FAC sort. We decided

390   to use the 293T cells as a proof-of-principle cell line and the PC3 cells as a relevant prostate

391   cancer model cell line, hence the HeLa cells were not included in studies from this point

392   onwards. PCR integration checks were performed on the 293T and PC3 clonal populations to

393   determine the efficiency of HITI- vs HDR-mediated reporter gene integration at the AAVS1 site

394   (Figure 2A-B). The number of 293T clonal populations with correct integration was 11.8%

395   (4/34) for HDR-AAVS1 engineered cells and 36.1% (13/36) for HITI-AAVS1 clones (Figure

396   2A-B). PC3 cells grew fewer colonies but showed zero integration at the AAVS1 site for tdT+

397   HDR engineered cells (0/14), whereas 10.5% (2/19) of the HITI engineered colonies had correct

398   reporter gene integration, indicating that HITI was more efficient in both cell types.

399

400   <u>*In vitro* reporter gene imaging</u>

401   Next, we expanded single 293T and PC3 clonal populations that had correct integration bands

402   for further *in vitro* reporter gene functionality testing. Firstly, we confirmed tdT fluorescence for

403   both the 293T-HITI (Figure 3A) and PC3-HITI (Figure 3G) clones via fluorescence microscopy.

404   In addition, there was a positive correlation between BLI signal and increasing cell number for

405   293T (Figure 3B-C; r^2=0.9718) and PC3 (Figure 3H-I; r^2=0.9897) cells. BLI signal measured

406   over several passages showed stable FLuc2 expression over time for both clonal cell lines

407   (Figure 3D and J). To test for Oatp1a1 functionality, 293T naïve, 293-HITI, PC3 naïve and PC3-

408   HITI cells were incubated with or without Gd-EOB-DTPA (6.4 mM) in normal media for 90

409   minutes, washed thoroughly, and pelleted before inserting into an agarose phantom. Inversion

410   recovery MRI was performed at 3 Tesla and 37°C, and spin-lattice relaxation rate (R_1) maps

411   were generated (Figure 3E, K). Neither the naïve 293T/PC3 or untreated 293T-HITI and PC3-

412   HITI cell populations exhibited any change in R_1 rates (Figure 3F, L). Only HITI clones

413   expressing Oatp1a1 had significantly increased R_1 rates after Gd-EOB-DTPA incubation, with

414   ~10-fold increase for 293T-HITI cells (7.952 ± 0.87 s^{-1}) compared with naïve, treated controls

415   (0.806 ± 0.038 s^{-1}; n = 3, $P < 0.001$; Figure 3F) and ~5-fold increase for PC3-HITI cells ($3.426 \pm$

416   0.217 s^{-1}) compared with naïve, treated controls (0.6402 ± 0.045 s^{-1}; n = 3, $P < 0.001$; Figure

417   3L).

418

419   <u>Oatp1a1 sensitivity *in vivo*</u>

420   To investigate the MR detection limit of Oatp1a1-expressing cells we injected various

421   combinations of PC3-naïve and PC3-HITI cells at five sites subcutaneously on the backs of nude

422   mice (Figure 4 and Supp. Figure 4). A total of $3x10^6$ cells were injected per site with the

423   following number of PC3-HITI cells: 1 – 0 (naïve cell only control); 2 – 10^4; 3 – 10^5; 4 – 10^6; 5 –

424    $3x10^6$ (HITI cell only control). Naïve cells were included with HITI cells so that each injection

425   contained a total of $3x10^6$/site. BLI signal intensity increased as PC3-HITI cell numbers

426   increased (representative mouse shown in Figure 4A), with 10^6 and $3x10^6$ HITI injections

427   showing significant signal increase above naïve background controls (Figure 4B). Transverse

428   MR images from the same mouse showed positive contrast at both the 10^6 and $3x10^6$ HITI

429   injection sites 5 hours after Gd-EOB-DTPA injection (Figure 4C). Similar to the BLI data, these

430   sites also exhibited significantly higher contrast-to-noise ratios (CNR) than naïve controls

431   (Figure 4D). The 10^4 and 10^5 PC3-HITI injections were difficult to visualize on MRI, had no

432   discernible positive contrast and were, therefore, not measured. These data were consistent

433   across all three mice (see Supp. Figure 4) and showed that the very minimum number of

434   Oatp1a1-expressing cells we could detect with Gd-EOB-DTPA-based MR contrast was 10^6 cells

435   in a 50 μl injection volume.

436

437   <u>PC3-HITI Oatp1a1 tumour models for MRI detection</u>

438   As a proof-of-principle that our HITI-engineered cells could show Gd-EOB-DTPA induced

439   positive MRI contrast in subcutaneous tumours, we injected 293T-naïve and 293T-HITI cells on

440   either flank of a nude mouse (Supp. Figure 5). For both cell types, the large masses were visible

441   on pre-contrast images but also showed noticeable positive contrast 20 minutes post-Gd-EOB-

442   DTPA injection. However, 5 hours post-contrast, the naïve tumour had returned to pre-contrast

443   background levels, whereas the HITI tumour had very prominent positive contrast that also

444   showed heterogeneity within the tumour mass (Supp. Figure 5), similarly to what we reported

445   previously[43].

446

447   Moving into a more relevant cancer model, we next injected PC3-naïve and PC3-HITI clonal

448   cells subcutaneously on either flank of nude mice and followed BLI development over time

449   (Figure 5). After only 11 days post-injection, before the tumours were visible or palpable, clear

450   positive contrast was observed for HITI engineered cells 5 hours after Gd-EOB-DTPA injection

451   (Figure 5A and Supp. Figure 6). The same mice were then imaged at day 46 where the naïve

452   tumour was visible due to pooled Gd-EOB-DTPA after 20 minutes post-contrast injection. In a

453   similar fashion 5 hours post contrast, only the HITI engineered cells retained the Gd agent and

454   showed bright, positive contrast (Figure 5B and Supp. Figure 7). Quantification of CNR over

455   time showed a significant increase for the HITI tumours (Figure 5C). These data suggest that the

456   Oatp1a1 MRI reporter can detect tumour burden at stages where the tumours are not visible or

457   palpable and tumour growth can be tracked longitudinally with Gd-EOB-DTPA-enhanced MRI.

458

459   Finally, PC3-HITI cells were injected via the tail vein into NSG mice to investigate the ability of

460   Oatp1a1 as a reporter gene for visualizing metastases (Figure 6). Using BLI as a guide, we were

461   able to detect a Gd-EOB-DTPA-enhanced metastatic tumour in the liver of one mouse (Figure

462   6A). In this case, it was evident that a shadow in the liver (as seen in the pre-contrast MR image)

463   was likely the tumour, which the post-contrast image confirmed. However, in another mouse that

464   had BLI signal in the head region, a cluster of smaller PC3-HITI tumours was only identifiable

465   in post-contrast images (Figure 6B), indicating the usefulness of this reporter gene for detecting

466   metastatic burden with contrast-enhanced MRI.

467

468   **Discussion**

469   As personalized medicine and CRISPR-editing become a reality in the clinic, there is a greater

470   need to 1) improve the efficiency, efficacy and safety of genetically engineered cell-therapies,

471   and 2) improve our understanding of disease progression and treatment response in preclinical

472   models of disease. Reporter gene-based imaging allows us to track the location, viability, growth

473   and efficacy of such treatments, and in preclinical models of cancer progression and treatment. In

474   this study, we have developed a non-viral vector-based engineering system for large DNA

475   multimodality reporter gene integration into the AAVS1 safe-harbour genomic locus. To

476   improve safety further, we utilized MCs as the DNA vector of choice, which eliminates bacterial

477   DNA contamination and antibiotic resistance genes. In addition, we showed that utilizing the

478   non-homologous end joining (NHEJ) repair pathway with HITI could improve DNA editing

479   efficiency in human cells compared to the more commonly used HDR pathway. Finally, building

480   off our previous work[32], we have engineered a trimodality reporter gene construct that contains a

481   clinically relevant MRI reporter, *Oatp1a1*, in addition to fluorescent and bioluminescent genes

482   which enabled cell sorting and non-invasive BLI/MRI of engineered cells in a pre-clinical cancer

483   model.

484

485   One of the major limitations of engineering cells with large, multimodality reporter gene DNA

486   plasmids is the reduced efficiency of both transfection and gene editing with increasing

487   construct/insert size[45–47]. In addition, the presence of bacterial and antibiotic resistance genes in

488   parental plasmids has the potential to exert immunological responses and raises safety concerns.

489   To circumvent these issues, we designed our study to use MCs, which removes the bacterial

490   backbone from parental plasmids and thus reduces the size of the DNA donor constructs. Using

491   MCs instead of parental plasmids allowed us to remove ~4 kb of unwanted DNA from our HDR

492   construct, with a further reduction of ~1.5 kb for the HITI MC when the homologous arm

493   sequences were replaced with a 20 bp gRNA sequence (saving a total of ~5.5 kb). These large-

494   scale reductions thus provided us with room to upgrade our dual-modality *tdTomato* and *FLuc2*

495   reporter gene construct we previously reported[32] to a trimodality reporter gene construct with the

496   addition of the *Oatp1a1* MRI reporter[42,43]. To improve safety and translatability we also removed

497   the puromycin resistance gene to reduce the MC size by a further 600 bp and utilized FACS of

498   tdTomato positive cells to obtain mixed and clonal cell populations instead of antibiotic

499   selection. Our final step for improving safety was to design our system to target a "safe harbour"

500   locus in genomic DNA. Several of these loci have now been reported in the literature[48] and are

501   described as sites where inserted genetic elements can function as intended, without causing

502   alterations that would pose a risk to the host cell or organism[23]. For this study we targeted the

503   AAVS1 site found within the human Protein Phosphatase 1 Regulatory Subunit 12C gene as this

504   has been one of the best characterized, to date. No known side effects are associated with

505   disrupting the PP1R12C gene, however it has been reported that mechanisms such as DNA

506   methylation can silence transgenes targeted to this genomic region[49]. Since our studies rely on

507   stable reporter gene expression over time for accurate cell detection and proliferation, we

508   investigated whether reporter gene expression in our AAVS1 engineered 293T-HITI and PC3-

509 HITI cell populations changed over time. We found that BLI signal was stable over several

510 passages and tdT fluorescence was consistently expressed in both cell lines, indicating consistent

511 transgene expression.

512

513 We have shown here that HITI-based CRISPR/Cas9 cell engineering is more efficient than the

514 more commonly used HDR method for integrating large DNA donor constructs into the genome

515 for stable expression. Targeted transgene integration is typically achieved using homologous

516 arms and the HDR pathway, however, this mechanism is highly inefficient and is not usually

517 active in non-dividing cells[33]. Indeed, our previous study showed only 3.8% of selected cells

518 were correctly edited using the HDR mechanism[32]. In contrast, the HITI method which utilizes

519 the NHEJ pathway is active in all stages of the cell cycle and in quiescent cells[36], and thus has

520 been used to improve editing efficiency. Using the method described by Suzuki *et al.*[38], our

521 engineered 293T and PC3 clonal cell populations did indeed have greater DNA integration at the

522 AAVS1 site compared with HDR (36% and 12% for HITI vs 10.5% and 0% for HDR,

523 respectively). However, it is important to note that the NHEJ repair pathway is error prone and

524 often leads to insertions and deletions (indels) at the DNA junctions. Consequently, this

525 mechanism is often taken advantage of to produce DNA disruptions, gene silencing and knock-

526 outs. These issues would need to be considered if using the HITI method for correctional DNA

527 editing and promoter-less vector integration, since these require specific DNA sequences, either

528 upstream or downstream, to be preserved. In this case, we engineered cells with the only

529 requirements being that the transgene inserts into the AAVS1 site (confirmed with junctional

530 PCR) and that the reporter genes are consistently expressed (confirmed with imaging).

531 Therefore, indels at either the 5' or 3' junction would likely have a negligible impact on our

532 experiments.

533

534 Although we confirmed correct transgene integration at the AAVS1 site in our study, we cannot

535 rule out integration at other off-target sites in HITI and HDR engineered populations. Several

536 293T and PC3 single-cell clonal populations expressed the *tdTomato* fluorescence reporter gene

537 but did not show integration bands for the AAVS1 site. Evidence suggests that CRISPR/Cas9 is

538 not 100% accurate and off-target effects have been reported[50,51], thus, it is possible that the MCs

539 integrated into off-target Cas9 cut sites. However, it is also plausible, and probably more likely,

540 that the MCs inserted into the AAVS1 site in the wrong direction. Even though HITI is designed

541 to minimize integration in the wrong orientation, the error-prone NHEJ repair mechanism of

542 blunt-ended DNA breaks could lead to indels at the CRISPR/Cas9 cut site boundaries, which

543 could then disrupt the ability of Cas9 to recognize and re-cut those sites. The likelihood of indel

544 formation using Cas9 HITI could be reduced in future studies by adopting a similar method to

545 that recently reported by Li and colleagues[52]. In that study, the authors utilize Cas12a, which

546 leaves 5 bp overhangs, to precisely edit the genome in a process coined microhomology-

547 dependent targeted integration (MITI)[52]. Independently of CRISPR, MCs, like plasmids, can also

548 randomly integrate into the genome of cells, albeit at very low rates. Future work will need to

549 analyze the rate of off-target integrations and possible indel disruptions at the CRISPR/Cas9 cut

550 sites using techniques, such as next generation sequencing, to determine the full safety profile of

551 HITI at safe harbour loci. To improve targeting specificity, studies have shown that high-fidelity

552 Cas9 enzymes in ribonucleoprotein complexes (RNPs), instead of Cas9 DNA vectors, improve

553 on-target activity, while reducing off-target editing[53,54]. In combination with RNPs, adeno-

554 associated viruses (AAVs) are now commonly used as DNA donors for CRISPR experiments

555 due to their high transduction capabilities in hard-to-transfect cell lines, their low risk of random

556 integration and reduced immunogenic response. However, AAVs are still limited by their

557 loading capacity of ~4.5 kb, which would be a problem for large, multimodality imaging vectors

558 as presented here, but conceivable for future studies where only one imaging reporter gene is

559 required. With these emerging technologies, it is likely that CRISPR gene editing will become

560 highly specific and thus safer in the near future.

561

562 We engineered cells with a multimodality reporter gene construct to enable us to go from single

563 cell, optical imaging methods (FLI) to higher sensitivity whole-animal planar imaging (BLI) and

564 superior 3D high-resolution tomographic imaging (MRI) in animals. This offers several

565 advantages. Firstly, fluorescently activated cell sorting of tdTomato-expressing cells improves

566 upon our previous study by eliminating the need for an antibiotic resistance selection gene,

567 which constitutes a safety risk and has been associated with structural plasmid instabilities[55].

568 Secondly, the firefly luciferase gene (*FLuc2*), in combination with its substrate D-luciferin,

569 allowed us to directly visualize engineered cells *in vivo* using BLI. Inclusion of bioluminescent

570 genes in preclinical cancer models is a relatively inexpensive and valuable tool that also allows

571   one to track cell migration and cell seeding in metastatic cancer models, assess cell viability and

572   follow cell/tumour growth longitudinally[14]. A limitation of BLI is that it is restricted to small

573   animal models of disease. However, it is useful for determining sites of cell

574   arrest/seeding/growth and thus can be used in conjunction with other reporter genes as a guide

575   for determining when and where to perform relatively more expensive, higher resolution clinical

576   imaging, such as MRI[13]. To build off our previous dual FLI-BLI study[32], we decided to include

577   the MRI reporter gene, *Oatp1a1*, as a translationally relevant and sensitive reporter gene to

578   complete our trimodality construct for HITI-based CRISPR engineering. First described by

579   Patrick *et al.*[42], *Oatp1a1* selectively, but reversibly, uptakes the clinically approved Gd^{3+} contrast

580   agent Gd-EOB-DTPA and thus provides positive contrast in T_1-weighted MR images. The

581   authors concluded, therefore, that *Oatp1a1* engineered cells and tumours should be easier to

582   detect than the negative contrast generated by T_2-agents, such as such as superparamagnetic iron

583   oxide (SPIO) and ferromagnetic agents[42,56]. In addition, engineering cells with integrated

584   *Oatp1a1* expression means that MR images can be obtained longitudinally to track cell migration

585   and growth, and signal intensity can be directly correlated with cell viability. Finally, we and

586   others have found that *Oatp1a1* also enhances the uptake of the firefly luciferase substrate D-

587   luciferin for BLI[43,57] and the fluorescent dye indocyanine green (using the human ortholog

588   *OATP1B3*) for both fluorescent[58] and photoacoustic imaging[59], which gives an added advantage

589   of using *Oatp1* for multi-modality imaging. Since we, and others, have now shown that the

590   human *OATP1B3* gene also functions as a useful fluorescent, photoacoustic and MRI reporter

591   gene, *in vivo*[58,59], future studies will focus on exchanging *Oatp1a1* for the more translationally

592   favourable *OATP1B3* ortholog.

593

594   The improved safety profile and expression of multimodal reporter genes proposed here could

595   have several uses in cell engineering, or at least help answer several concerns with *in vivo* cell

596   therapies. For example, the U.S. Food and Drug Administration have listed potential safety

597   concerns related to unproven stem cell therapies[60], including: 1) the ability of cells to move from

598   placement sites and change into inappropriate cell types or multiply; 2) failure of cells to work as

599   expected, and 3) the growth of tumours. In addition, the long-term safety profiles of cells

600   engineered with randomly-integrating viruses still require further investigation and optimization.

601   These are concerns that could be addressed by targeting non-viral DNA vectors, such as MCs, to

602   specific safe-harbour loci, such as AAVS1, and reducing the use of integrating viruses.

603   Incorporating reporter genes for clinical grade imaging will also help improve patient safety by

604   allowing one to track cellular therapies *in vivo* (such as for stem cells and immunotherapies).

605   Clinicians could then determine whether the therapeutic cells are localizing to the correct

606   anatomical feature, such as a solid tumour[18], or to determine their persistence and viability for

607   short- and long-term treatment strategies. Future work will focus on evaluating our system in

608   stem cells and other clinically relevant cell types. Translation will also need to consider building

609   donor vectors that lack optical reporter genes and utilize other selection methods (e.g., magnetic

610   sorting). It is easily feasible to switch out genes from our trimodality construct for other imaging

611   purposes, such as replacing *FLuc2* with a PET reporter gene for dual PET-MR imaging. Suicide

612   switch genes could also be incorporated to further improve safety by killing the engineered cells

613   in cases where they become oncogenic[61], for example. Not only are these tools useful for clinical

614   cell-based therapies, they are also extremely useful in pre-clinical studies for investigating cancer

615   progression/aggression, metastatic burden and treatment strategies. Avoiding the use of random-

616   integrating viruses and targeted editing should also help reduce off-target effects of gene-editing

617   that may alter the normal characteristics of the cell type being studied.

618

619   Conclusion

620   Our work demonstrates the first CRISPR/Cas9 HITI MC system for safe harbor integration of a

621   large donor construct encoding three reporter genes for multi-modal longitudinal imaging of cells

622   *in vivo*. We have shown that inclusion of the translationally relevant MR reporter gene, *Oatp1a1*,

623   can enable localization and tracking of small primary and metastatic tumours that are not readily

624   detectable visually or in pre-contrast MR images. This work lays the foundation for an effective

625   and safer genome editing tool for non-invasive reporter gene tracking of multiple cell types *in*

626   *vivo,* such as for cell-based cancer immunotherapies and stem cell treatments.

627

628   Author Contributions

629   J.A.R. designed the project, J.A.R, J.J.K. and M.S.M. designed the experiments. J.J.K. directed

630   the study and with M.S.M. carried out most of the experiments. N.N.N. developed the methods

631   for Oatp1a1 MRI. Y.C. helped perform MRI. M.M.E. analyzed MRI data. A.J.H. helped develop

632   the parental plasmids. J.J.K. wrote the manuscript with help from M.S.M., which J.A.R.

633   reviewed and edited.

634

Acknowledgements

639

Competing Interests

641   The authors declare that they have no competing interests.

642

Funding

648

References

1. Brader, P., Serganova, I. & Blasberg, R. G. Noninvasive molecular imaging using reporter genes. *Journal of Nuclear Medicine* **54**, 167–172 (2013).

2. Kircher, M. F., Gambhir, S. S. & Grimm, J. Noninvasive cell-tracking methods. *Nat. Rev. Clin. Oncol.* **8**, 677–688 (2011).

3. Prescher, J. A. & Contag, C. H. Guided by the light: visualizing biomolecular processes in living animals with bioluminescence. *Current Opinion in Chemical Biology* **14**, 80–89 (2010).

4. Hong, H., Yang, Y. & Cai, W. Imaging gene expression in live cells and tissues. *Cold Spring Harb. Protoc.* **6**, (2011).

5. Kim, J. E., Kalimuthu, S. & Ahn, B. C. In Vivo Cell Tracking with Bioluminescence Imaging. *Nucl. Med. Mol. Imaging (2010).* **49**, 3–10 (2015).

6. Li, M., Wang, Y., Liu, M. & Lan, X. Multimodality reporter gene imaging: Construction strategies and application. *Theranostics* **8**, 2954–2973 (2018).

7. Gilad, A. A. & Shapiro, M. G. Molecular Imaging in Synthetic Biology, and Synthetic Biology in Molecular Imaging. *Mol. Imaging Biol.* **19**, 373–378 (2017).

8. Joo, H. K. & Chung, J. K. Molecular-genetic imaging based on reporter gene expression. *J. Nucl. Med.* **49**, 164–180 (2008).

9. Reagan, M. R. & Kaplan, D. L. Concise Review: Mesenchymal Stem Cell Tumor-Homing: Detection Methods in Disease Model Systems. *Stem Cells* **29**, 920–927 (2011).

10. Wang, H. *et al.* Trafficking Mesenchymal Stem Cell Engraftment and Differentiation in Tumor-Bearing Mice by Bioluminescence Imaging. *Stem Cells* **27**, 1548–1558 (2009).

11. Kidd, S. *et al.* Direct evidence of mesenchymal stem cell tropism for tumor and wounding microenvironments using in vivo bioluminescent imaging. *Stem Cells* **27**, 2614–2623 (2009).

12. Hamilton, A. M., Parkins, K. M., Murrell, D. H., Ronald, J. A. & Foster, P. J. Investigating the Impact of a Primary Tumor on Metastasis and Dormancy Using MRI: New Insights into the Mechanism of Concomitant Tumor Resistance. *Tomography* **2**, 79–84 (2016).

13. Parkins, K. M. *et al.* Multimodality cellular and molecular imaging of concomitant tumour enhancement in a syngeneic mouse model of breast cancer metastasis. *Sci. Rep.* **8**, (2018).

681   14.   Parkins, K. M. *et al.* A multimodality imaging model to track viable breast cancer cells

682   from single arrest to metastasis in the mouse brain. *Sci. Rep.* **6**, (2016).

683   15.   Vandergaast, R. *et al.* Enhanced noninvasive imaging of oncology models using the NIS

684   reporter gene and bioluminescence imaging. *Cancer Gene Therapy* (2019).

685   doi:10.1038/s41417-019-0081-2

686   16.   Shah, K., Jacobs, A., Breakefield, X. O. & Weissleder, R. Molecular imaging of gene

687   therapy for cancer. *Gene Therapy* **11**, 1175–1187 (2004).

688   17.   Yaghoubi, S. S. *et al.* Noninvasive detection of therapeutic cytolytic T cells with 18 F-

689   FHBG PET in a patient with glioma. *Nat. Clin. Pract. Oncol.* **6**, 53–58 (2009).

690   18.   Keu, K. V. *et al.* Reporter gene imaging of targeted T cell immunotherapy in recurrent

691   glioma. *Sci. Transl. Med.* **9**, (2017).

692   19.   Milone, M. C. & O'Doherty, U. Clinical use of lentiviral vectors. *Leukemia* **32**, 1529–

693   1541 (2018).

694   20.   Hacein-Bey-Abina, S. *et al.* Insertional oncogenesis in 4 patients after retrovirus-mediated

695   gene therapy of SCID-X1. *J. Clin. Invest.* **118**, 3132–3142 (2008).

696   21.   Howe, S. J. *et al.* Insertional mutagenesis combined with acquired somatic mutations

697   causes leukemogenesis following gene therapy of SCID-X1 patients. *J. Clin. Invest.* **118**,

698   3143–3150 (2008).

699   22.   Hacein-Bey-Abina, S. *et al.* LMO2-Associated Clonal T Cell Proliferation in Two Patients

700   after Gene Therapy for SCID-X1. *Science (80-.).* **302**, 415–419 (2003).

701   23.   Papapetrou, E. P. & Schambach, A. Gene Insertion Into Genomic Safe Harbors for Human

702   Gene Therapy. *Mol. Ther.* **24**, 678–684 (2016).

703   24.   Wang, Y. *et al.* Genome editing of human embryonic stem cells and induced pluripotent

704   stem cells with zinc finger nucleases for cellular imaging. *Circ. Res.* **111**, 1494–1503

705   (2012).

706   25.   Luo, Y. *et al.* Stable Enhanced Green Fluorescent Protein Expression After Differentiation

707   and Transplantation of Reporter Human Induced Pluripotent Stem Cells Generated by

708   AAVS1 Transcription Activator-Like Effector Nucleases. *Stem Cells Transl. Med.* **3**, 821–

709   835 (2014).

710   26.   Cerbini, T. *et al.* Transcription activator-like effector nuclease (TALEN)-mediated

711   CLYBL targeting enables enhanced transgene expression and one-step generation of dual

712   reporter human induced pluripotent stem cell (iPSC) and neural stem cell (NSC) lines.

713   *PLoS One* **10**, (2015).

714   27.   Cho, S. W., Kim, S., Kim, J. M. & Kim, J. S. Targeted genome engineering in human cells

715   with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **31**, 230–232 (2013).

716   28.   Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science (80-.).*

717   **339**, 819–823 (2013).

718   29.   Jinek, M. *et al.* RNA-programmed genome editing in human cells. *Elife* **2013**, (2013).

719   30.   Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science (80-.).* **339**,

720   823–826 (2013).

721   31.   Ding, Q. *et al.* Enhanced efficiency of human pluripotent stem cell genome editing

722   through replacing TALENs with CRISPRs. *Cell Stem Cell* **12**, 393–394 (2013).

723   32.   Dubois, V. P. *et al.* Safe Harbor Targeted CRISPR-Cas9 Tools for Molecular-Genetic

724   Imaging of Cells in Living Subjects. *Cris. J.* **1**, 440–449 (2018).

725   33.   Orthwein, A. *et al.* A mechanism for the suppression of homologous recombination in G1

726   cells. *Nature* **528**, 422–426 (2015).

727   34.   Lieber, M. R. The Mechanism of Double-Strand DNA Break Repair by the

728   Nonhomologous DNA End-Joining Pathway. *Annu. Rev. Biochem.* **79**, 181–211 (2010).

729   35.   Auer, T. O., Duroure, K., De Cian, A., Concordet, J. P. & Del Bene, F. Highly efficient

730   CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair.

731   *Genome Res.* **24**, 142–153 (2014).

732   36.   Suzuki, K. & Izpisua Belmonte, J. C. In vivo genome editing via the HITI method as a

733   tool for gene therapy. *J. Hum. Genet.* **63**, 157–164 (2018).

734   37.   He, X. *et al.* Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced

735   homology-dependent and independent DNA repair. *Nucleic Acids Res.* **44**, (2016).

736   38.   Suzuki, K. *et al.* In vivo genome editing via CRISPR/Cas9 mediated homology-

737   independent targeted integration. *Nature* **540**, 144–149 (2016).

738   39.   Wang, T. D., Chen, Y. & Ronald, J. A. A novel approach for assessment of prostate

739   cancer aggressiveness using survivin-driven tumour-activatable minicircles. *Gene Ther.*

740   **26**, 177–186 (2019).

741   40.   Ronald, J. A. *et al.* Development and Validation of Non-Integrative, Self-Limited, and

742   Replicating Minicircles for Safe Reporter Gene Imaging of Cell-Based Therapies. *PLoS*

743   *One* **8**, (2013).

744   41.   Darquet, A. M., Cameron, B., Wils, P., Scherman, D. & Crouzet, J. A new DNA vehicle

745   for nonviral gene delivery: Supercoiled minicircle. *Gene Ther.* **4**, 1341–1349 (1997).

746   42.   Patrick, P. S. *et al.* Dual-modality gene reporter for in vivo imaging. *Proc. Natl. Acad. Sci.*

747   *U. S. A.* **111**, 415–20 (2014).

748   43.   Nyström, N. N. *et al.* Longitudinal Visualization of Viable Cancer Cell Intratumoral

749   Distribution in Mouse Models Using Oatp1a1 -Enhanced Magnetic Resonance Imaging.

750   *Invest. Radiol.* **54**, 302–311 (2019).

751   44.   Yushkevich, P. A. *et al.* User-guided 3D active contour segmentation of anatomical

752   structures: Significantly improved efficiency and reliability. *Neuroimage* **31**, 1116–1128

753   (2006).

754   45.   Paix, A. *et al.* Precision genome editing using synthesis-dependent repair of Cas9-induced

755   DNA breaks. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E10745–E10754 (2017).

756   46.   Kreiss, P. *et al.* Plasmid DNA size does not affect the physicochemical properties of

757   lipoplexes but modulates gene transfer efficiency. *Nucleic Acids Res.* **27**, 3792–3798

758   (1999).

759   47.   Hornstein, B. D., Roman, D., Arévalo-Soliz, L. M., Engevik, M. A. & Zechiedrich, L.

760   Effects of circular DNA length on transfection efficiency by electroporation into HeLa

761   cells. *PLoS One* **11**, (2016).

762   48.   Pellenz, S. *et al.* New Human Chromosomal Sites with 'Safe Harbor' Potential for

763   Targeted Transgene Insertion. *Hum. Gene Ther.* **30**, 814–828 (2019).

764   49.   Ordovás, L. *et al.* Efficient recombinase-mediated cassette exchange in hPSCs to study the

765   hepatocyte lineage reveals AAVS1 locus-mediated transgene inhibition. *Stem Cell Reports*

766   **5**, 918–931 (2015).

767   50.   Zhang, X. H., Tee, L. Y., Wang, X. G., Huang, Q. S. & Yang, S. H. Off-target effects in

768   CRISPR/Cas9-mediated genome engineering. *Molecular Therapy - Nucleic Acids* **4**, e264

769   (2015).

770   51.   Wang, D. C. & Wang, X. Off-target genome editing: A new discipline of gene science and

771   a new class of medicine. *Cell Biology and Toxicology* **35**, 179–183 (2019).

772   52.   Li, P. *et al.* Cas12a mediates efficient and precise endogenous gene tagging via MITI:

773   microhomology-dependent targeted integrations. *Cell. Mol. Life Sci.* (2019).

774   doi:10.1007/s00018-019-03396-8

775   53.   Vakulskas, C. A. *et al.* A high-fidelity Cas9 mutant delivered as a ribonucleoprotein

776   complex enables efficient gene editing in human hematopoietic stem and progenitor cells.

777   *Nat. Med.* **24**, 1216–1224 (2018).

778   54.   Chen, J. S. *et al.* Enhanced proofreading governs CRISPR-Cas9 targeting accuracy.

779   *Nature* **550**, 407–410 (2017).

780   55.   Oliveira, P. H. & Mairhofer, J. Marker-free plasmids for biotechnological applications -

781   implications and perspectives. *Trends in Biotechnology* **31**, 539–547 (2013).

782   56.   Xiao, Y. D. *et al.* MRI contrast agents: Classification and application (Review).

783   *International Journal of Molecular Medicine* **38**, 1319–1326 (2016).

784   57.   Patrick, P. S., Lyons, S. K., Rodrigues, T. B. & Brindle, K. M. Oatp1 Enhances

785   Bioluminescence by Acting as a Plasma Membrane Transporter for d-luciferin. *Mol.*

786   *Imaging Biol.* **16**, 626–634 (2014).

787   58.   Wu, M. R. *et al.* Organic anion-transporting polypeptide 1B3 as a dual reporter gene for

788   fluorescence and magnetic resonance imaging. *FASEB J.* **32**, 1705–1715 (2018).

789   59.   Nyström, N. N., Yip, L. C. M., Carson, J. J. L., Scholl, T. J. & Ronald, J. A. Development

790   of a Human Photoacoustic Imaging Reporter Gene Using the Clinical Dye Indocyanine

791   Green. *Radiol. Imaging Cancer* **1**, e190035 (2019).

792   60.   Marks, P. W., Witten, C. M. & Califf, R. M. Clarifying Stem-Cell Therapy's Benefits and

793   Risks. *N. Engl. J. Med.* **376**, 1007–1009 (2017).

794   61.   Ivics, Z. Self-Destruct Genetic Switch to Safeguard iPS Cells. *Molecular Therapy* **23**,

795   1417–1420 (2015).

796

Figure 1
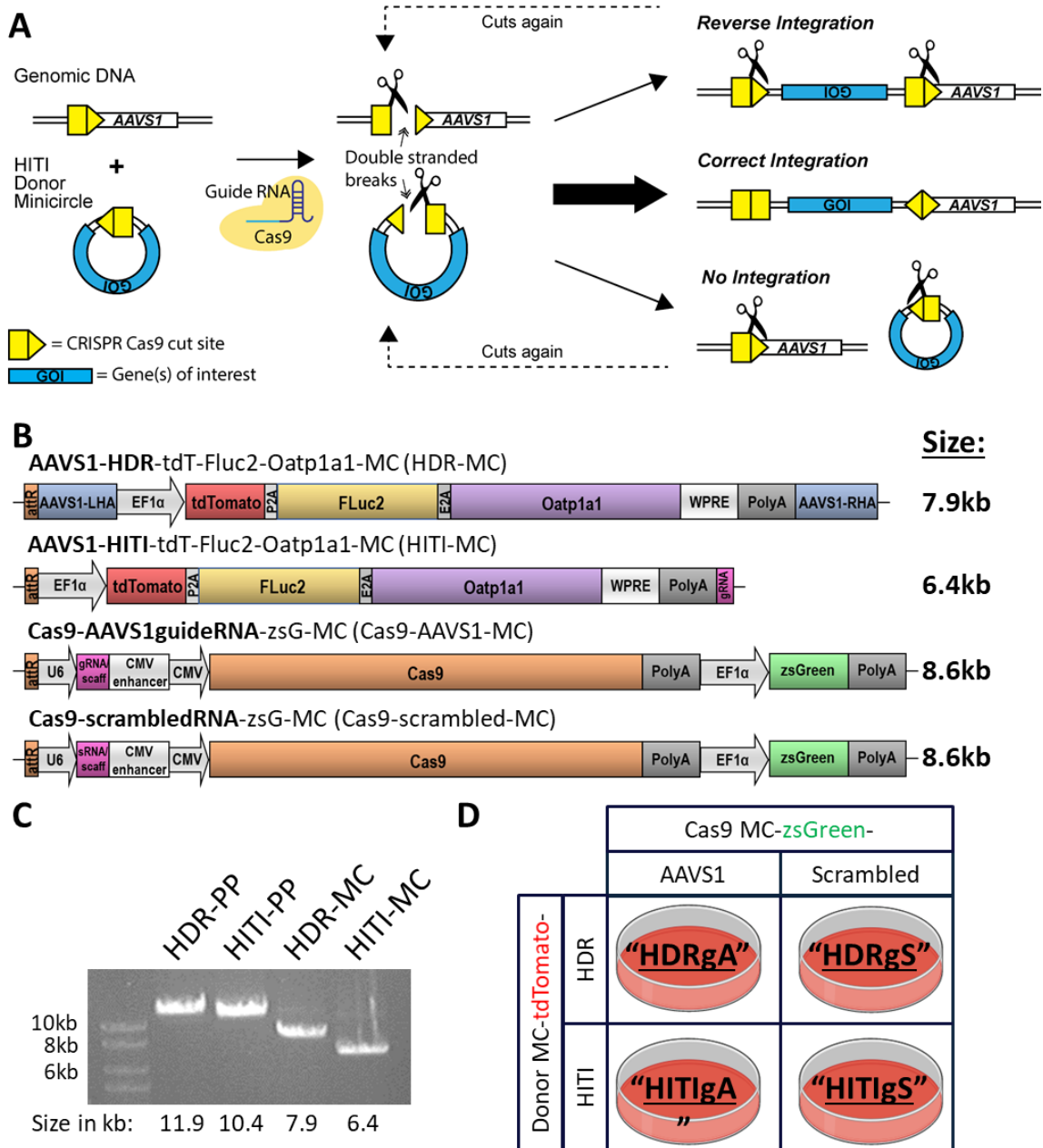
Figure 1. Homology-Independent Targeted Integration (HITI) experimental design. *A*, HITI minicircles (MCs) contain a Cas9 cut site identical to that at the AAVS1 safe harbor locus. Both genomic and MC DNA are cut in the presence of a guide RNA and Cas9. Genes of interest (GOI) are only stably integrated into the genome when inserted in the correct orientation, otherwise the Cas9 cut sites are preserved, which increases the likelihood of continuous Cas9 cutting. Image adapted from Suzuki *et al.*, 2016. *B*, Trimodality HDR, HITI and Cas9 MC constructs designed for this study. *C*, Restriction digest agarose gel of parental plasmids (PP) and MCs and indicated band sizes. *D*, transfection regimen for combinations of donor and Cas9 MCs and simplified abbreviations for each condition.

Figure 2

A
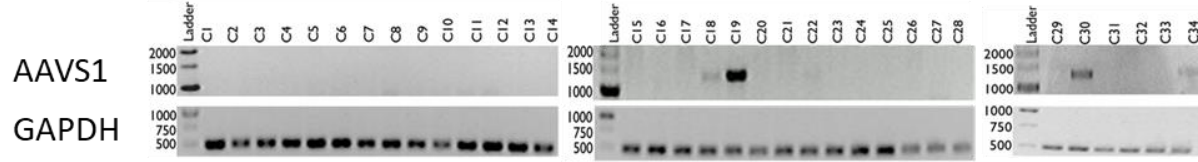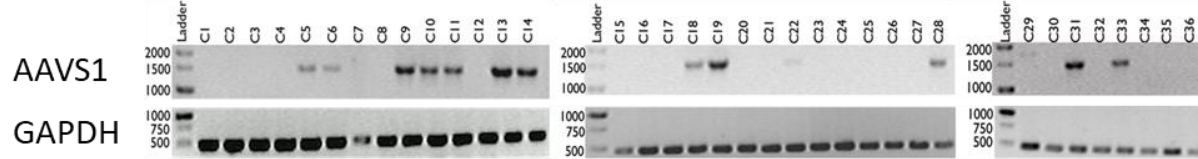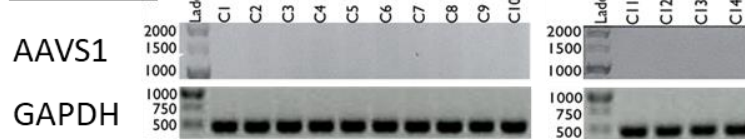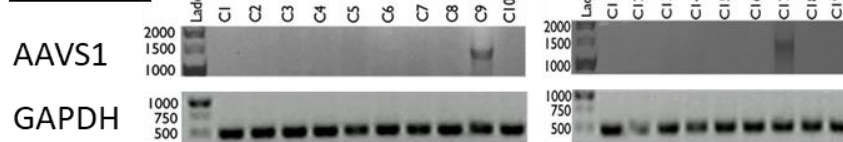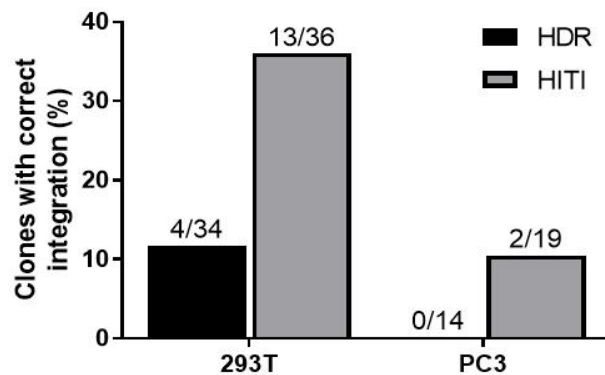

Figure 2. Junctional PCR integration checks for 293T and PC3 clonal cell populations. *A*, PCR integration checks at the AAVS1 site. GAPDH was amplified as a DNA loading control. *B*, Quantification shows a higher number of positive integration clones for HITI engineered cells compared to HDR for both 293T and PC3 cell lines.

Figure 3
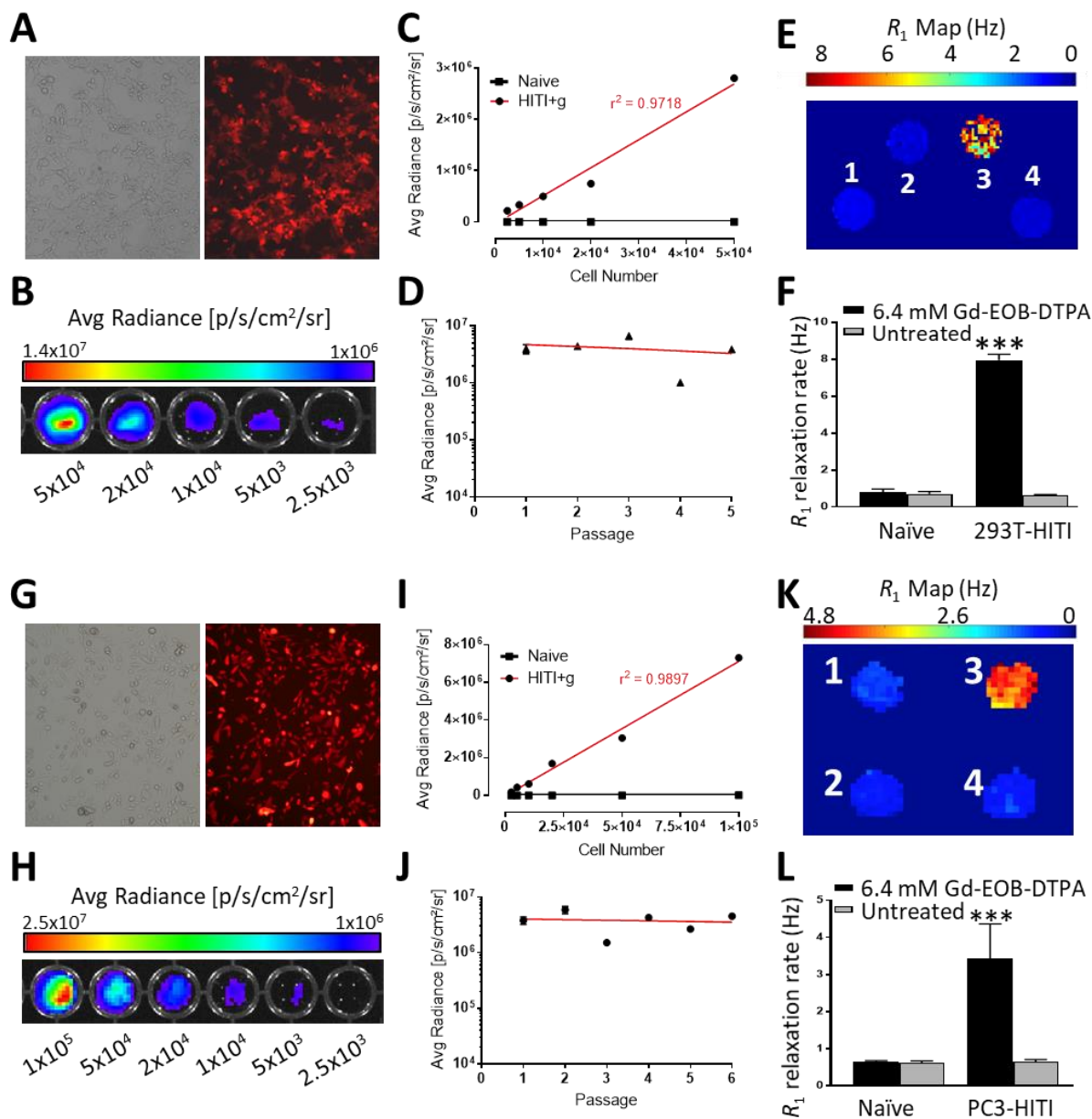
Figure 3. *In vitro* FLI, BLI and MRI characterization of 293T-HITI (*A-F*) and PC3-HITI (*G-L*) clones. *A* and *G*, brightfield and tdT fluorescence. *B* and *H*, BLI intensity maps related to cell number. *C* and *I*, quantification of BLI signal to cell number. *D* and *J*, BLI signal over successive passages. *E* and *K*, Spin-lattice relaxation maps of representative phantoms containing pellets of cells untreated or treated with 6.4 mM Gd-EOB-DTPA, as follows: 1, naive, treated; 2, naïve untreated; 3, HITI treated; 4, HITI untreated. *F* and *L*, quantification of spin-lattice rates (n = 3, *** $P < 0.001$).
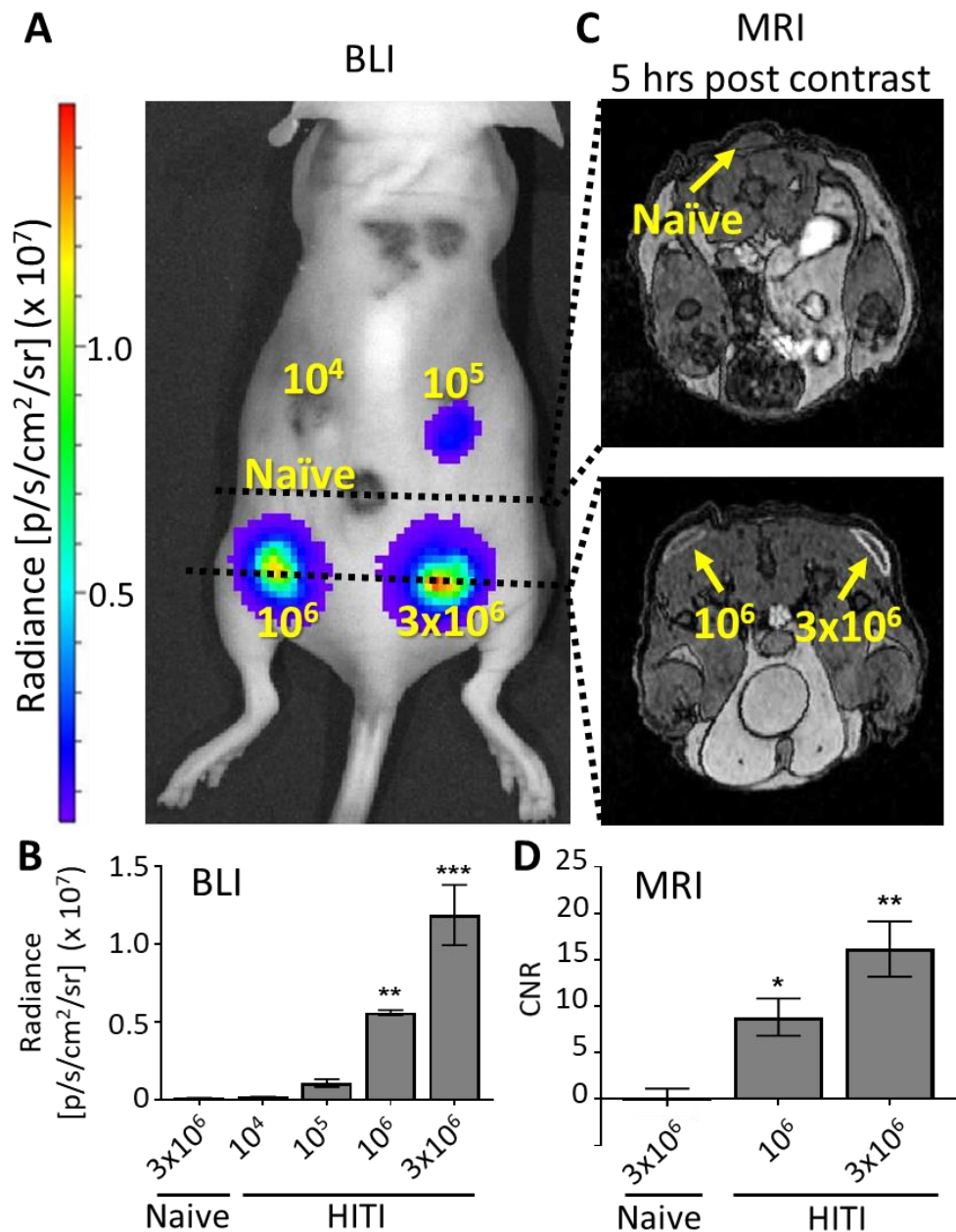
Figure 4. BLI and Oatp1a1 sensitivity *in vivo*. *A*, A total of 3×10^6 PC3 cells were injected subcutaneously into five locations on the back of nude mice, with increasing concentrations of HITI engineered cells, as indicated in yellow, and corresponding BLI signals. *B*, quantification of BLI signal from ROIs around the five PC3 injection sites. *C*, MRI transverse views of cell injection sites 5 hours after Gd-EOB-DTPA injection. *D*, quantification of contrast-to-noise (CNR) ratio for 10^6 and 3×10^6 PC3-HITI cells. Note, 10^4 and 10^5 PC3-HITI injections lacked enough contrast to measure CNR values. n = 3 mice, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.
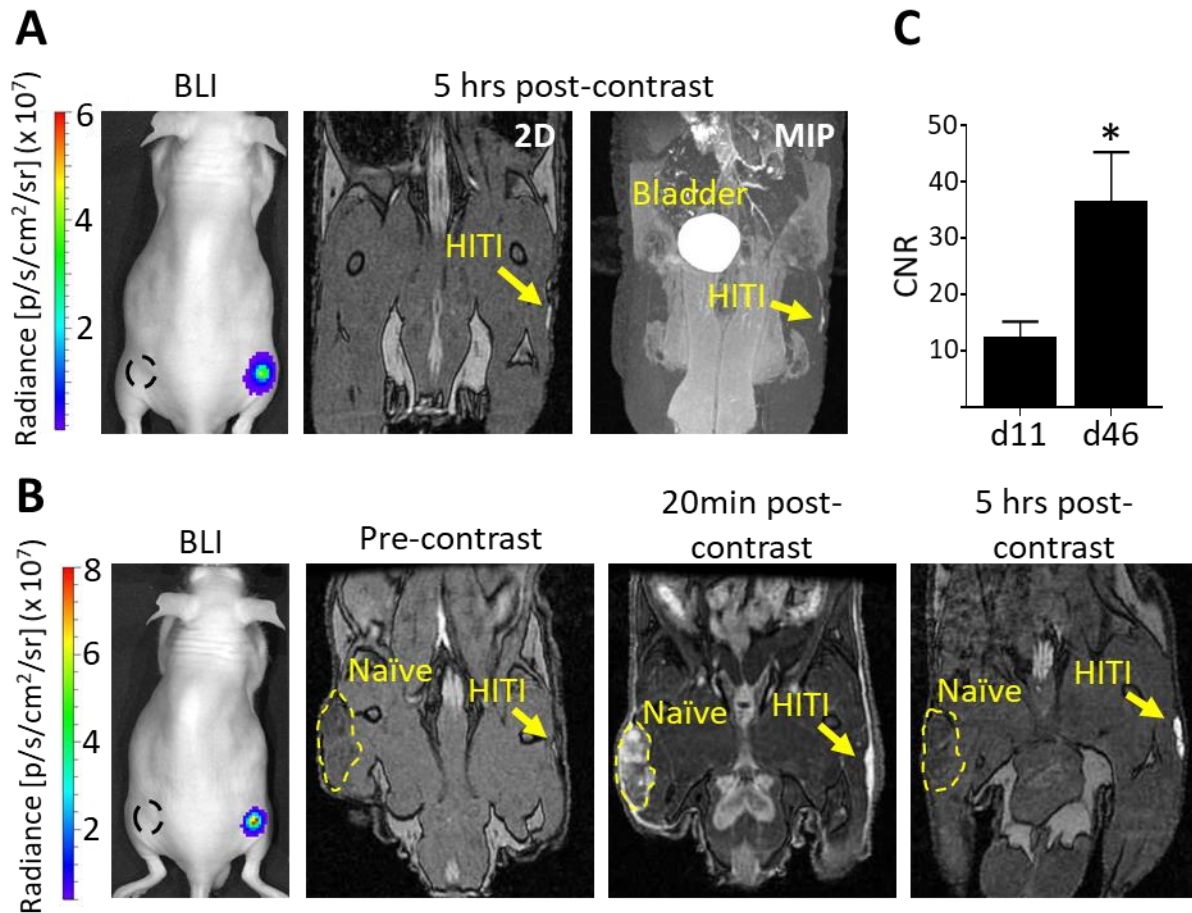
Figure 5. Longitudinal *in vivo* MRI of subcutaneous PC3-HITI cells. Mice were injected subcutaneously with 1×10^6 naïve and PC3-HITI cells on the left and right flanks, respectively. BLI signal was present on right flank only. Naïve tumour locations are denoted by black dashed line. *A*, day 11 post PC3 injection. 2D and maximum intensity projection (MIP) images acquired 5 hrs post Gd-EOB-DTPA injection. *B*, the same mouse was re-imaged at day 46. Pre-, 20min post- and 5hr post-contrast coronal images were obtained. *C*, contrast-to-noise ratios (CNR) of PC3-HITI tumours 5hr post-contrast showed significant increase from day 11 to day 46. n = 4 (day 11) and n = 3 (day 46) * *P* < 0.05.

Figure 6

A

B
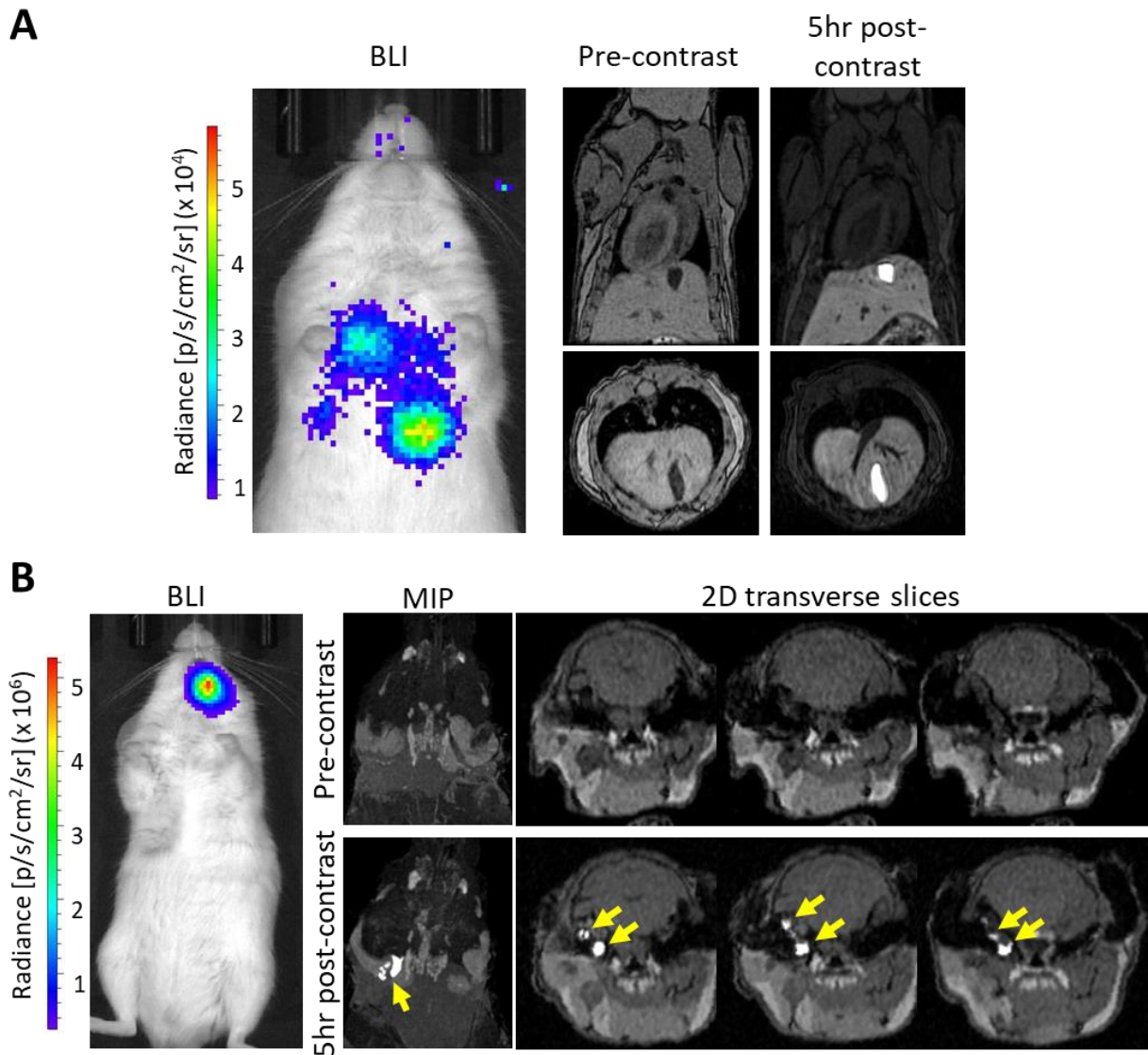
Figure 6. *In vivo* MRI detection of metastatic PC3-HITI tumours. *A*, The brightest BLI signal corresponded to a tumour in the liver, which is evident as a shadow in pre-contrast images and as a bright positive Gd-EOB-DTPA contrast tumour in images taken 5 hrs post-contrast injection. *B*, A mouse showed BLI signal in the head region, which was not evident in pre-contrast MIP and 2D transverse MRI slices (upper panel). However, clusters of PC3-HITI tumours were easily discernible in 5hr post-contrast images (yellow arrows, lower panel).