

# A caveat in mouse genetic engineering: ectopic gene targeting in ES cells by bidirectional extension of the homology arms of a gene replacement vector carrying human *PARP-1*

Aswin Mangerich · Harry Scherthan · Jörg Diefenbach ·  
Ulrich Kloz · Franciscus van der Hoeven ·  
Sascha Beneke · Alexander Bürkle

**Abstract** Here we report an approach to generate a knock-in mouse model using an ‘ends-out’ gene replacement vector to substitute the murine *Parp-1* (*mParp-1*) coding sequence (32 kb) with its human orthologous sequence (46 kb). Unexpectedly, examination of mutant ES cell clones and mice revealed that site-specific homologous recombination was mimicked in three independently generated ES cell clones by bidirectional extension of the vector homology arms using the endogenous *mParp-1*-flanking sequences as templates. This was followed by adjacent integration of the targeting vector, thus leaving the endogenous *mParp-1* locus functional. A related phenomenon termed ‘ectopic gene targeting’ has so far only been described for ‘ends-in’ integration-type vectors in

non-ES cell gene targeting. We provide reliable techniques to detect such ectopic gene targeting which represents an unexpected caveat in mouse genetic engineering that should be considered in the design and validation strategy of future gene knock-in approaches.

**Keywords** ES cells · Gene targeting · Homologous recombination · Knock-in mice · *PARP-1*

## Abbreviations

DAPI	4',6-Diamidino-2-phenylindol-dihydrochlorid
DSB	Double strand break
DTA	Diphtheria toxin A
ESC	Embryonic stem cell clone
F	Filial generation
FISH	Fluorescence in situ hybridization
hPARP-1	Human PARP-1
HR	Homologous recombination
mParp-1	Murine Parp-1
Neo <sup>R</sup>	Neomycin resistance
PARP-1	Poly(ADP-ribose) polymerase-1
qPCR	Quantitative PCR
SDSA	Synthesis-dependent strand annealing
wt	Wild-type

A. Mangerich · J. Diefenbach · S. Beneke ·  
A. Bürkle (✉)  
Molecular Toxicology Group, Department of Biology,  
University of Konstanz, Box X911, 78457 Constance,  
Germany  
e-mail: alexander.buerkle@uni-konstanz.de

H. Scherthan  
Bundeswehr Institute of Radiobiology, Munich, Germany

U. Kloz · F. van der Hoeven  
German Cancer Research Center, Transgenic Core  
Facility, Heidelberg, Germany

## Introduction

More than 20 years after its first description, gene targeting by homologous recombination (HR) in

murine embryonic stem (ES) cells has become a standard method in biomedical research and has provided a wealth of valuable data to study mammalian gene function on a cellular and organismal level (Capecchi 2005; Mak 2007). However, gene targeting efficacy and outcome remain variable, for reasons that are not fully understood (Glaser et al. 2005). Mechanistically, the technology relies on HR as an error-free cellular repair process for DNA double strand breaks (DSB) (Helleday et al. 2007; Li and Heyer 2008). Upon detection of a DSB, the DNA 5' end is resected by an as yet unknown nuclease and the remaining single-stranded 3' overhang is coated by Rad51. The resulting Rad51 nucleofilament then invades the homologous double-stranded DNA (mainly sister chromatids), followed by annealing with the complementary DNA strand (Sun et al. 1991; Barzel and Kupiec 2008). The resulting D-loop structure serves as a priming end for DNA synthesis using the intact DNA strand as a template. At the invasion site an X-shaped structure is then formed known as the Holliday junction. This junction can slide in either direction of the crossing point in a process known as branch migration. Dependent on the direction of the sliding process two outcomes, both resulting in the repair of the DNA break, are possible. First, if the sliding occurs in the same direction as the DNA synthesis, the invading strand can dissociate and re-anneal with the other end of the trimmed DSB, thereby restoring a functional error-free DNA molecule. This synthesis-dependent strand annealing (SDSA) pathway manages restoration of proper DNA structure without employing crossing-over events and is thought to be the predominant mechanism during HR (Helleday et al. 2007). Second, if branch migration takes place in the opposite direction accompanied by strand invasion of the other end of the DSB, this can result in the prominent double Holliday junction structure. The newly synthesized DNA strand can then reattach to the other end of the DNA break. Subsequent cleavage and relegation of the DNA strands can lead to a crossing-over event, i.e., resolution of the double Holliday junction by exchange of the double-stranded DNA arms, and finally repair of the DSB (Helleday et al. 2007; Barzel and Kupiec 2008).

Two basic types of vectors can be distinguished in classical gene targeting, i.e., 'ends-in' integration and 'ends-out' replacement vectors (van der Weyden

et al. 2002). Integration-type vectors are linearized within the homology region, resulting in homology arms pointing towards each other and final integration of the vector at the locus of interest in a single crossing over event. In the case of replacement-type vectors, the sequence to be introduced is flanked by two homology arms and the vector is linearized within the vector backbone outside the homology regions (van der Weyden et al. 2002). Li and Baker suggested that the two homologous arms of a gene replacement vector invade the chromosome at the target locus. Complete exchange of the sequences located in between the homology is then caused by resolution of the Holliday structure by crossing-over events in each flanking arm (Li and Baker 2000; Li et al. 2001). Although SDSA is thought to be the predominant mechanism of DSB repair (Helleday et al. 2007), its role in gene targeting is not well understood.

By using the technique of gene targeting a total of about 9,000 genes have so far been knocked out in the mouse genome, with deletion of sequences up to 30 kb (Zhang et al. 1994; Brault et al. 2006; Collins et al. 2007). In recent years, application of gene targeting has been expanded to the generation of so called gene knock-in mice, which now become increasingly interesting for biomedical research (Roebroek et al. 2003). In this approach the gene of interest is replaced with a different functional sequence under the transcriptional control of *cis*- and *trans*-acting elements that are part of the endogenous gene (Babinet and Cohen-Tannoudji 2001). The field of application for this method ranges from exchange of the gene of interest with a reporter gene (Elefanty et al. 1998; Luche et al. 2007) to the generation of humanized mice (Rozmahel et al. 1997; Luo et al. 2001), mainly by replacing loci with sequences of up to 10 kb, such as cDNA. However, the replacement of larger sequences might be interesting as well, e.g., the exchange of whole coding sequences including all introns to study the function of the introduced gene under "physiological" transcriptional control. For this purpose, recombination-based methods using chromosome engineering can be exploited, as shown recently with the generation of a  $\alpha$ -globin humanized mouse model (Wallace et al. 2007). Nevertheless, studies like these are elaborate and costly as they require at least two successive rounds of gene targeting.

Here we report an approach using classical gene targeting with an ‘ends-out’ gene-replacement vector to generate a knock-in mouse model in order to substitute the murine *Parp-1* coding sequence (32 kb) with its human orthologous sequence (46 kb). Unexpectedly, despite the bona fide correct gene targeting, as identified by long-range flanking polymerase chain reactions (PCRs) and Southern blot analyses, our approach resulted in ectopic integration of the targeting vector. Yet, site-specific HR was mimicked by bidirectional extension of the vector homology arms, presumably by employing a SDSA-like pathway. Our results provide proof-of-principle that a phenomenon which was formerly described as ‘ectopic gene targeting’ when using ‘ends-in’ integration-type vectors during non-ES cell gene targeting, can also occur during gene targeting with ‘ends-out’ gene-replacement vectors in mouse ES cells. Thereby, this report demonstrates a potential caveat in mouse genetic engineering and gives new insight into the mechanisms that can take place during mammalian gene targeting. Moreover, we provide reliable techniques to identify such obstacles, which might be useful in particular with respect to the generation of large knock-in genomic alterations introduced by traditional gene targeting.

## Methods

### Cloning of the *hPARP-1* targeting vector and gene targeting in ES cells

The targeting construct was cloned based on a pBR322 vector backbone by performing four classical cloning steps with two fusion-PCR fragments and two cloning steps employing HR in *E. coli* DY380 (Copeland et al. 2001; Lee et al. 2001). Human *PARP-1* DNA was obtained from human BAC RP11-15H13 harboring the entire *hPARP-1* coding sequence. Mouse homologous regions of 5 kb each at the 5' and 3' end of the *mParp-1* coding sequence were obtained by proofreading PCR from mouse genomic DNA isolated from E14-1 ES cells. For positive and negative selection of ES cell clones, a “floxed” *Neo<sup>R</sup>* cassette was included in intron 14 and a diphtheria toxin A (*DTA*) was fused to the 3' homologous region. All exons and exon/intron borders were sequenced. No mutations were detected

except for a 50-bp deletion in a non-conserved region of intron 1.

### Gene targeting and generation of chimeras

Gene targeting and generation of chimeras was performed using standard procedures (Nagy 2003). Briefly, before electroporation of ES cells the targeting vector was linearized by *Pfl23II* digestion. Successful linearization was examined by control digestion with *XmaIII* and *NotI*, respectively, and evaluation of the resulting DNA fragment fingerprints. Murine ES cells (line E14-1 derived from strain 129P2/OlaHsd) were grown on feeder cells and electroporated with 50 µg of the *Pfl23II*-linearized targeting vector, followed by selection with 330 µg/ml G418. After 9 days of selection, 276 ES cell clones were picked. One aliquot of each clone was expanded for DNA extraction and another aliquot cryoconserved for blastocyst injection. Clones were screened by using a flanking PCR method and results were verified by Southern blotting. Bona fide positive clones were thawed, expanded, and used for generation of chimeras. For this purpose, donor females were superovulated with gonadotropin (i.p.) and 47 h later with chorionic gonadotropin, and then mated with fertile males. Successfully mated females were identified by plug check and sacrificed by cervical dislocation. Embryos at morula stage (day 2.5 of embryonic development) were isolated from these mice, washed through several droplets of medium and incubated overnight in a CO<sub>2</sub> incubator. The blastocysts that had developed, were injected with 10–15 genetically modified ES cells of clones #113, #225, or #267. The injected blastocysts were incubated for 1–2 h in a CO<sub>2</sub> incubator and thereafter, transferred into the uteri of pseudopregnant recipients, which gave birth to chimeric offspring.

### Genetic characterization by PCR and Southern-blot analysis

General genotyping for *mParp-1* and *hPARP-1* was performed with a PCR reaction using a mixture of three primers: the *mParp-1* wild-type allele was detected with the *mParp-1*-specific sense primer 5'-CTT GAG AAG AAT GCT CTG GGC TCT GAA G-3' located in intron 20 resulting in an amplicon of 220 bp, the *hPARP-1* coding sequence

was detected with the *hPARP-1*-specific sense primer 5'-GCT TGA TTT GGA CAT GGA CCC TAG TCC T-3' located in intron 20 resulting in an amplicon of 421 bp. The antisense primer 5'-TTG CCC TTG GGT AAC TTG CTG ATA TGT G-3' binds to both *PARP-1* orthologous sequences and is located in exon 21. Genotyping for the neomycin-resistance cassette was performed with the sense primer 5'-TGC TCC TGC CGA GAA AGT ATC CAT CAT GGC-3' and the antisense primer 5'-CGC CAA GCT CTT CAG CAA TAT CAC GGG TAG-3' resulting in an amplicon of 380 bp. As an internal PCR control the murine *Fabpi* gene was amplified using the sense primer 5'-TGG ACA GGA CTG GAC CTC TGC TTT CCT AGA-3' and the antisense primer 5'-TAG AGC TTT GCC ACA TCA CAG GTC ATT CAG-3' resulting in an amplicon of 194 bp. Absence of the *DTA* cassette in the targeted genome was examined using the sense primer 5'-GTG GTC AAA GTG ACG TAT CCA GGA CTG ACG-3' and the antisense primer 5'-AGC ACT ACA CGC GAA GCA CCA TCA TCG AAC-3' resulting in an amplicon of 170 bp. As an internal PCR control, primers for the *Neo<sup>R</sup>* cassette were used. PCR reactions were incubated at 94°C for 2 min and then 34 cycles of amplification at 94°C for 20 s, 62°C for 20 s, and 68°C for 1 min with a terminal elongation for 10 min at 68°C in a Biorad thermal cycler PTC-200.

Site-specific integration of the targeting construct and screening of the targeted ES cell clones was examined by long-range PCR reactions amplifying regions spanning the homology arms at the promoter and at the terminator side of *Parp-1*. In the case of the promoter flanking PCR the sense primers 5'-GGA GCC TGG TCC TTC GGG TAC C-3' or 5'-TGC CACCACGCCAGCTTGC-3' bind upstream the homologous regions in the mouse genome and the human-*PARP-1*-specific antisense primer 5'-CGC CCG CAC AGC GGC CC-3' binds in intron 1 resulting in the case of bona fide correct gene targeting in amplicons of 5.2 kb or 7.2 kb, respectively. In the case of the terminator flanking PCR the *hPARP-1*-specific sense primer 5'-TGA GGC ACA TGT ACA TAC CC-3' binds in intron 22 of the *hPARP-1* coding sequence and the antisense primers 5'-TCC ACT TGT TAG GGG ACC C-3' or 5'-TAT GAAGCACAGAGCAGTCC-3' bind downstream of the homologous regions in the mouse genome resulting in the case of bona fide correct gene

targeting in amplicons of 5.3 kb or 8 kb, respectively. Hot-start PCR reactions were incubated at 94°C for 3 min and then 36 cycles of amplification at 94°C for 10 s, 68°C (promoter PCR) or 62°C (terminator PCR) for 20 s, and 68°C for 6–8 min with an terminal elongation of 10 min at 68°C in a Biorad thermal cycler PTC-200.

Amplicons were separated by 0.8–2% agarose gel electrophoresis and detected by ethidium bromide staining. Flanking PCR amplicons were sequenced by GATC Biotech to rule out artefacts and results were confirmed by Southern blot analyses.

Southern analysis was performed according to standard techniques (Sambrook 2001) by digestion of 10–15 µg of genomic DNA with *KpnI* or *HindIII* (promoter side, 'internal' or 'external' probes, respectively) or *BsrGI* (terminator side) for at least 14 h, separated by agarose gel electrophoreses for about 5 h, transferred onto a Hybond-N+ membrane (Amersham) by capillary wet blot technique for at least 16 h, and hybridized with <sup>32</sup>P-radio-labeled probes as shown in Fig. 2a. To reduce background signal, salmon sperm DNA was added in a concentration of 100 µg/ml to the hybridization buffer of membranes that were subsequently hybridized with the promoter 'external' probe. Autoradiograms were exposed for 7 days.

#### Chromosome preparation and fluorescence in situ hybridization

For metaphase preparation ES cells were treated with BrdU (45 µM) for 12 h and with 10 µg/ml Colcemid for 1 h at 37°C before trypsin harvesting. For hypotonic shock, 10 ml of a prewarmed (37°C) 75 mM KCl solution was slowly added to less than  $1 \times 10^7$  cells, followed by incubation for 20 min at 37°C. Thereafter, 2–3 drops of ice-cold fixative (acetic acid/methanol 1:3) were added and cells pelleted for 5 min at 250g. Using vortexing, 10 ml of ice-cold fixative was added drop-wise to the pellet. Samples were pelleted and the procedure repeated three times. Chromosomes were stored in fixative at –20°C until further use.

Human or murine *Parp-1* probes were generated based on the targeting construct or murine E14-1 genomic DNA, respectively, using proofreading PCR amplification of 18-kb weakly-conserved intron-rich regions. Initial amplicons were reamplified by the degenerative oligonucleotide primer (DOP) PCR

amplification kit (Roche) according to the manufacturer's instructions.

Labeling and hybridization of FISH-probes was performed as described by Scherthan et al. (1994). The human probe was labeled with biotin-dUTP (Bionick nick translation kit, Invitrogen), murine probe with digoxigenin-dUTP (dig nick kit, Roche) by nick translation. Biotin was detected by fluorescein-conjugated avidin (ExtrAvidin-FITC, Sigma), digoxigenin using an anti-digoxigenin-antibody conjugated with tetrahydroaminoisothiocyanate (TRITC, Roche). Hybridized spreads were evaluated using a Zeiss epifluorescence microscope. Microphotographs were recorded using the ISIS imaging system (Meta-Systems).

Determination of gene copy numbers  
by quantitative real-time PCR

Mouse genomic DNA was prepared using the High Pure DNA Template Purification kit (Roche). Human *PARP-1*, *Gm821*, and *Lin9* gene copy numbers were determined by Sybr-green real-time PCR approach using the Biorad iCycler system. Sense primer sequences were 5'-ATCTGAGCTCCATGGACAGG-3', 5'-CTTCCTGCTTCTCGTTGGTC, GCCTCCCTGGAGAAAATAGG-3' and antisense primer sequences 5'-ACCCAGGCACAACAATACC-3', 5'-TTGCATGAGTAGGCTGGTTG-3', 5'-GACAAGATGGCCAGGAAGAG-3', respectively, amplifying products of 97–99 bp. A highly conserved sequence of *Cygb* with identical primer binding sites and amplicon sizes for human and murine sequences was used as a DNA standard (sense primer 5'-CAACACTGTCGTGAGAACC-3', antisense primer 5'-GGTTCACCTGTGCTTGAG-3').

PCRs for all genes were performed with 100 ng of DNA per sample in triplicates. Reactions were incubated at 95°C for 90 s and then 50 cycles of amplification at 95°C for 15 min and at 66°C for 1 min. Thereafter, melting curves of amplified products were determined, which confirmed specificity of the reactions.

Murine *Parp-1* gene copy numbers were determined using a Taq-Man approach with the *mParp-1*-specific probe FAM-AGGTGACAGAGAGCGGCCAGTGC-BHQ in combination with sense primer 5'-TCCAGAAAGCAGCGCC-3' and antisense primer 5'-GCACCCACAGCTGTGAACT-3' amplifying

a 71 bp product of *mParp-1* exon 8. For DNA standardization the actin probe FAM-ACTGCCG CATCCTTCTCCTCCC-BHQ1 was used in combination with sense primer 5'-TGCCTGACATCAAAGAGAAG-3' and antisense primer 5'-CAGCTCATAGCTCTTCTCC-3'. PCRs for both genes were performed with 100 ng of DNA per sample in triplicates. Reactions were incubated at 95°C for 1 min and then 40 cycles of amplification at 95°C for 15 s, at 55°C for 20 s, and at 68°C for 1 min. Data evaluation was performed using the  $\Delta\Delta C_t$  method.

RT-PCR analysis of PARP-1 mRNA expression

Total RNA was isolated using an RNeasy kit (Qiagen). Double-stranded cDNA was synthesized using the Omniscript RT kit (Qiagen). Murine or human cDNA was detected with the sense primers 5'-TGGC ACTGCCCTCTCTGTC-3' (murine exon 8), 5'-ACGCCTCCGCCCTCCACA-3' (human exon 8) and the antisense primers 5'-GGCCTTGTGGCAGATCCTGT-3' (murine exon 9), 5'-TCCTGGAGGAAGTCC TCAGACACA-3' (human exon 10). PCR reactions were incubated at 94°C for 2 min and then 36 cycles of amplification at 94°C for 15 s, 62°C for 20 s, and 68°C for 30 s with a terminal elongation of 10 min at 68°C in a Biorad thermal cycler PTC-200. Amplicon sizes are 179 and 281 bp in the case of the murine and human PARP-1 cDNA, respectively.

Western-blot analysis

Spleen tissue, snap-frozen in liquid nitrogen, was lysed on ice by homogenizing tissue with a douncer in modified high-salt RIPA buffer (9 ml/g tissue) containing 50 mM Tris-HCl (pH 7.4), 500 mM NaCl, 1% (w/v) Triton-X-100, 0.1% SDS, 1% (w/v) sodium-deoxycholate, and Complete protease inhibitor mix (Roche). Then, insoluble organ and cell debris was pelleted for 15 min at 4,000g and 4°C, protein concentrations were determined using the BCA protein assay (Pierce), diluted in 1.5× high-urea lysis buffer containing 62.5 mM Tris-HCl, pH 6.8, 6 M urea, 5% (v/v) 2-mercaptoethanol, 10% (v/v) glycerol, 2% (w/v) SDS, and 0.003% bromophenol blue. Forty micrograms of protein were used for western-blot analyses.

Whole-tissue lysates were separated on 8% polyacrylamide gels and proteins were transferred on Hybond-ECL nitrocellulose membrane (Amersham)

for 2 h at 300 mA using a wet blotting approach. Membranes were blocked for 1 h in 5% (w/v) milk powder in 10 mM Tris-HCl, pH 8.0, 150 mM NaCl, and 0.05% (v/v) Tween 20, and incubated overnight at 4°C with the following primary antibodies: monoclonal CII10 (PARP-1-specific), monoclonal FI23 (human-PARP-1-specific), and polyclonal antiserum 4595 (murine-PARP-1-specific). The murine-PARP-1-specific antibody 4595 was raised against a 14-aa peptide (SAPAPLALPLSVTS) of the mPARP-1 showing no homology with its human orthologue. Peroxidase-coupled goat-anti-mouse or goat-anti-rabbit secondary antibodies were used for chemiluminescent detection.

### Mouse breeding and housing

Experiments had been approved by the local veterinary authorities (Regierungspräsidium Freiburg, Germany). Mice were housed in the central animal care facility of the University of Konstanz, Germany, in isolated, ventilated cages (IVC) according to federal guidelines. Animals were killed by CO<sub>2</sub> inhalation.

### Statistics

Statistical analyses were performed with *Chi*-square test as indicated using GraphPad Prism software. *P* values  $\leq 0.05$  were considered statistically significant.

## Results

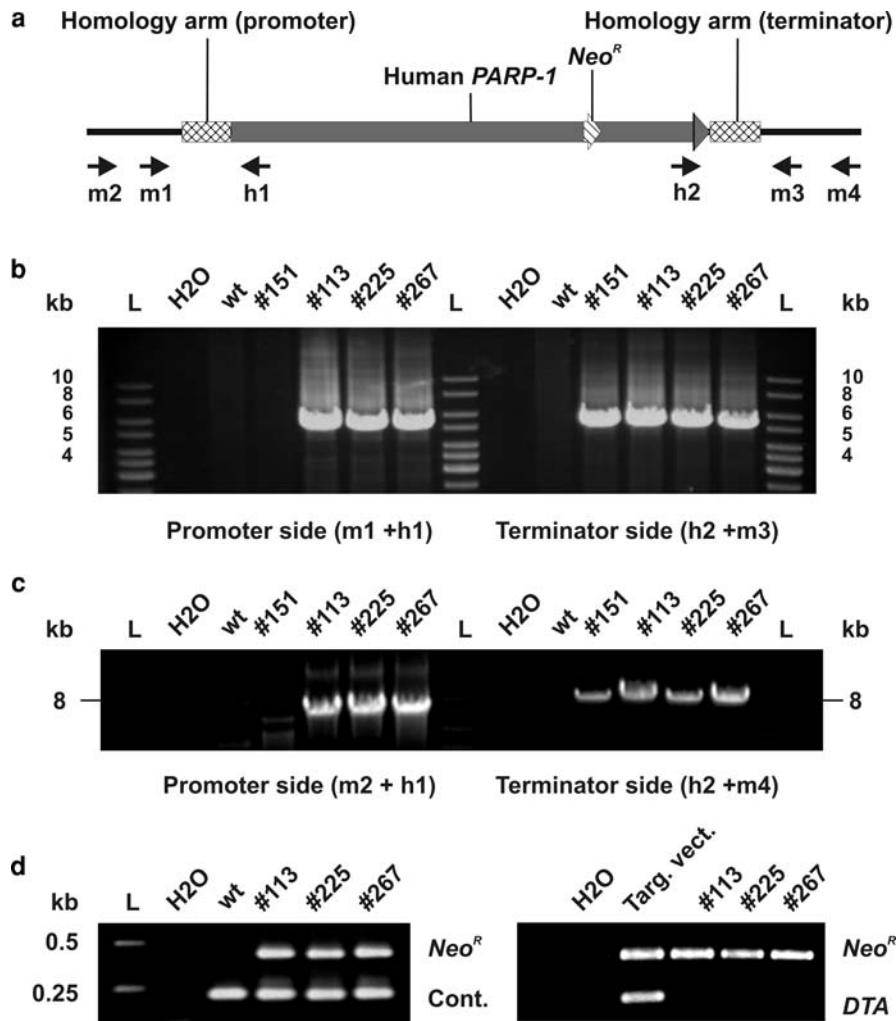
### Gene targeting strategy

We attempted to generate a human poly(ADP-ribose) polymerase-1 (*hPARP-1*) (GenBank ID: 142) knock-in mouse model by replacing the murine *Parp-1* (*mParp-1*) (GenBank ID: 11545) coding sequence (32 kb) with its human orthologous sequence (46 kb). For this purpose, an ‘ends-out’ gene-replacement targeting vector of 64 kb (Fig. S1 in Electronic supplementary material) was constructed using the technique of recombineering in *E. coli* (Copeland et al. 2001; Lee et al. 2001). The targeting vector comprised the entire *hPARP-1* coding sequence including all 23 exons and 22 introns flanked by 5 kb of murine regulatory sequences (i.e., promoter

and terminator) to permit site-specific HR in the mouse genome. A neomycin resistance (*Neo<sup>R</sup>*) cassette was included in intron 14 and a *DTA* cassette was placed adjacent to the homologous terminator sequences to enable positive–negative selection of ES cell clones (Yagi et al. 1990; van der Weyden et al. 2002). The *Neo<sup>R</sup>* cassette was flanked by two *loxP* sites to enable its Cre-mediated excision. Prior to electroporation, the targeting vector was linearized within the vector backbone by digestion with the restriction enzyme *Pfl23II* in order to generate the ‘ends-out’ configuration of the homology arms. Correct DNA-linearization was proven by the DNA fragment pattern obtained after control digestion with two different additional restriction enzymes (Fig. S2).

### PCR-based identification of ES cell clones with site-specific gene targeting

Upon electroporation of murine ES cells, 276 G418-resistant ES cell clones were picked and screened for site-specific HR performing four different long-range PCRs (Fig. 1a). Judging from (1) the specificity of the PCRs for site-specific *hPARP-1* knock-in as proven by the absence of amplicons when using murine wild-type DNA and (2) correct amplicon sizes of about 5 and 8 kb, three independently generated ES cell clones (#113, #225, #267; picked from different cell culture plates) displayed site-specific HR (Fig. 1b, c). This corresponded to a targeting efficiency of 1.1%. DNA sequencing of the obtained flanking PCR amplicons confirmed bona fide site-specificity of the targeting vector location, since promoter as well as terminator amplicons revealed the expected non-homologous murine sequences at their distal ends (sequencing with primers m1, m2, m3, and m4) and *hPARP-1*-specific sequences at their proximal ends (sequencing with primers h1 and h2) (Fig. S3). When using a mixture of wild-type ES cell genomic DNA and targeting vector DNA as templates, after PCR no formation of amplicons was observed (data not shown). Therefore, the hypothetical possibility of formation of the specific amplicons by internal fusion-PCR through the homology arms of the targeting vector and the corresponding regions of the endogenous *mParp-1* locus could be ruled out. Consistent with the general gene targeting strategy and the flanking PCR results, the presence of the



**Fig. 1** Identification of bona fide site-specific targeted ES cells by PCR. **a** Using a flanking PCR strategy, G418-resistant ES cell clones were screened for site-specific homologous recombination. In each PCR one primer bound specifically to the human sequence and its partner primer to the adjacent non-homologous sequences within the mouse genome, thereby flanking the homology arms of the targeting vector. For both homology arms (i.e., promoter and terminator) two different

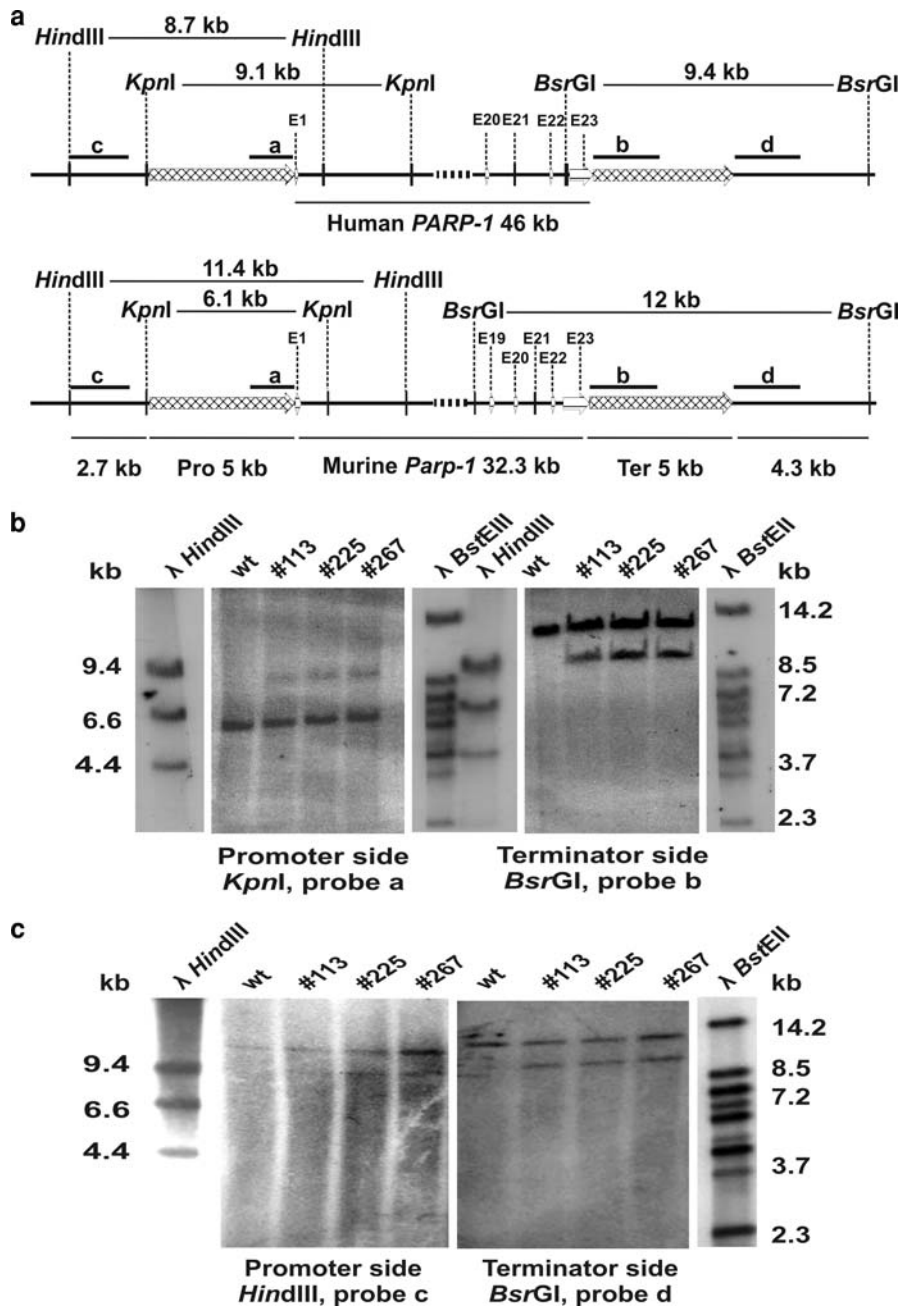
PCRs were performed resulting in amplicons of about **b** 5 kb or **c** 8 kb, respectively. Three clones were identified to be positive for site-specific homologous recombination, #113, #225, and #267. Clone #151 showed site-specific recombination only at the terminator homology arm. **d** Presence of the *Neo<sup>R</sup>* cassette and loss of the *DTA* cassette in the clones identified positive was verified by conventional genotyping PCR. *Cont.*, control; *L*, molecular size ladder; *Targ. vect.*, targeting vector

*Neo<sup>R</sup>* cassette and loss of the *DTA* cassette was confirmed in clones #113, #225, and #267 (Fig. 1d).

Verification of bona fide site-specific HR by Southern blot analyses

Southern blot analyses of genomic DNA isolated from G418-resistant ES cell clones are considered as the gold standard for the detection of site-specific HR

of the targeting vector (Nagy 2003). To verify the PCR results Southern analyses were performed using 'internal' probes (probes a and b) that bind within the vector homology arms, as well as 'external' probes (probes c and d) that either bind 5' (promoter) or 3' (terminator) of the homology arms only within the flanking mouse genome sequence (Fig. 2a). All fragment sizes expected for site-specific knock-in of the *hPARP-1* sequence were recovered although



**Fig. 2** Southern blot analyses of genomic DNA from targeted ES cells. **a** Southern blot strategy. DNA from ES cell clones, #113, #225, or #267, which were found positive for site-specific homologous recombination during PCR screening, was digested with restriction enzymes and sequentially hybridized with  $^{32}\text{P}$ -labeled probes as indicated. The indicated restriction fragments span the promoter (*pro*) or terminator (*ter*) homology arms. Internal probes, i.e., probes *a* (1.5 kb) and *b* (2 kb), bind within the vector homology arms. External probes, i.e., probes *c* and *d* (each 2 kb), bind 5' or 3' of the homology arms, respectively, within the adjacent non-homologous mouse genome. **b**

Hybridization of *KpnI* and *BsrGI*-digested genomic DNA with 'internal' probes for promoter and terminator fragments, respectively. **c** Hybridization of *HindIII* and *BsrGI*-digested genomic DNA with 'external' probes for promoter and terminator fragments, respectively. Note the weaker band intensities of the *hPARP-1* knock-in fragments compared to the fragments of *mParp-1* loci in the heterozygous mutant ES cell clones #113, #225, #267. *HindIII* and *BsrEII*-digested  $\lambda$ -genomic DNA served as a molecular weight standard. *E* indicates exon

signal intensities of the *hPARP-1* knock-in fragments were lower than those of the wild-type fragments. Intensities of promoter probes (9.1 kb, ‘internal’ probe a; 8.7 kb, ‘external’ probe c) were at the detection limit in the case of the *hPARP-1* knock-in fragments (Fig. 2b, c). (NB: An additional blot with mouse genomic DNA, which is comparable to the terminator blot of Fig 2c, is shown in Fig. 5b.)

#### Generation of chimeras and germline transmission

Based on the results from four independent PCR and Southern blot analyses, respectively, demonstrating bona fide site-specific HR of the *hPARP-1* targeting construct in the genomes of ES cell clones #113, #225, and #276, cells of these clones were injected into blastocysts from C57BL/6 donor mice. After transfer of chimeric blastocysts into uteri of foster mothers, we obtained 22 high-grade chimeras (chimerism higher than 50%, F0). The mutant mice were designated “B6;129P2 TgH (*hPARP-1*)”, hereafter referred to as “*hPARP-1* mice”. Chimeras were crossed with C57BL/6 females to obtain fully heterozygous *hPARP-1* knock-in mice. Germline transmission of the complete *hPARP-1* coding sequence was detected both by the flanking PCR strategy mentioned above

(see Fig. S4) and Southern blotting (see Fig. 5b) in 3 females of line #113 and 12 females and 9 males of line #225, respectively (F1). Germline transmission was not observed in the case of the single chimera of line #267. The observed ratio of mutant to wild-type was not significantly different from the expected ratio of 1:1 (*Chi*-square test  $P = 0.15$ ; 35 wild-type animals vs. 24 mutant animals).

#### Breeding of *hPARP-1* mice

Human *PARP-1* heterozygous mice of line #225 (F1) were intercrossed to generate homozygous *hPARP-1* knock-in mice. In concordance with Mendelian distribution, offspring comprised one quarter of wild-type and three quarters of mutant mice (Table 1a). Unexpectedly, none of the mutant animals lost *mParp-1*, which would have been expected in the case of the *hPARP-1* homozygosity. Selective intrauterine death of *hPARP-1* homozygotes was unlikely because the difference of the observed ratio of wild-type to mutant animals (1:3) compared to the ratio expected in the case of potential prenatal death of the homozygotes (1:2) showed ‘borderline’ significance (Table 1a). In addition, genotyping of embryos at day 12.5 of embryonic development revealed nearly the same genotypic ratio as found for

**Table 1** Genotype distribution, as analyzed by conventional PCR (line #225 F2) and qPCR (lines #225 and #113), of offspring from interbred heterozygous *hPARP-1* mice

(a) Genotype distribution, as analyzed by conventional PCR					
Genotype	Offspring (neonatal)	Expected (neonatal)	Expected in case of prenatal death of (h/h)	Offspring (E 12.5 dpc)	Expected (E 12.5 dpc)
Wild-type (m/m)	53 (27%)	25%	33%	8 (24%)	25%
<i>hPARP-1</i> (h/m)	141 (73%)	50%	67%	26 (76%)	50%
<i>hPARP-1</i> (w/o m)	0 (0%)	25%	0%	0 (0%)	25%
<i>P</i> value		<0.0001	0.09 (n.s.)		<0.01
(b) Genotype distribution, as analyzed by qPCR					
Genotype	Offspring #113	Offspring #225	Expected		
Wild-type (m/m)	9 (27%)	17 (35%)	25%		
Heterozygotes (1× <i>hPARP-1</i> )	17 (52%)	19 (42%)	50%		
Homozygotes (2× <i>hPARP-1</i> )	7 (21%)	12 (27%)	25%		
<i>P</i> value	0.87	0.21			

(a) Neonatal: 28 litters, mean litter size = 7. Prenatal: 6 litters, mean litter size = 10.2 embryos

(b) Four litters of line #113 and 6 litters of line #225 analyzed, mean litter sizes 8.25 and 8, respectively. Statistical analysis was performed by *Chi* square test. 1×*hPARP-1*, heterozygous mutant mice; 2×*hPARP-1*, homozygous mutant mice; *dpc*, days post coitum; *h*, human; *m*, murine

neonatal offspring (1:3) and once again did not identify any animals with loss of *mParp-1*. Intercrossing of mutant heterozygotes of line #113 (F2) did not result in any offspring with loss of *mParp-1* either.

Heterozygous *hPARP-1* mice (F1, both lines #113 and #225) were crossed with homozygous transgenic mice expressing Cre-recombinase ubiquitously (B6.FVB Tg (*EIIa-Cre*) C5379LMGD/JF) to excise the *Neo<sup>R</sup>* cassette. Offspring (F2) with excised *Neo<sup>R</sup>* was mated with C57BL/6 wild-type mice to outcross the *Cre* cassette (data not shown). The resulting heterozygous mutant mice (F3, lines #113 and #225) were intercrossed by brother–sister mating to obtain cohorts of wild-type as well as mutant heterozygous and homozygous animals (F4) lacking the *Neo<sup>R</sup>* and the *Cre* cassette. Backcrossing in a C57BL/6 background was continued for subsequent generations, and at the time of this writing two different *hPARP-1* congenic lines (F5) are available that are currently undergoing detailed phenotypic analysis.

#### Genomic localization of *hPARP-1* and *mParp-1* by fluorescence in situ hybridization

To search for the hypothetical presence of a remaining *mParp-1* sequence at the targeted locus on mouse chromosome 1, fluorescence in situ hybridization (FISH) was performed. Two different hybridization probe sets for *hPARP-1* and *mParp-1*, respectively, were generated by proofreading PCR followed by PCR amplification using the DOP approach. The human-specific probe spanned an 18-kb region from intron 1 to intron 4 of *hPARP-1* and the murine-specific probe an 18-kb region from intron 1 to intron 9 of *mParp-1*. Two color FISH experiments on mouse E14-1 ES cells and human A549 cells showed only hybridization on the homologous species' chromosomes demonstrating the specificity of the probes (Fig. 3a–c). Analyses of ES cell metaphase spreads excluded the presence of an additional chromosome 1, i.e., trisomy 1. FISH analyses on metaphase spreads of ES cell clones #113 and #225 demonstrated the integration of *hPARP-1* on chromosome 1 in both clones (Fig. 3d, e). Consistent with the breeding data, FISH analyses revealed that a copy of the endogenous *mParp-1* gene was remaining at the targeted locus (Fig. 3d, e). Due to the close proximity of the two signals (kb up to several Mb), it

was not possible to determine if the integration site of the targeting construct was distal or proximal of the endogenous *mParp-1* locus on chromosome 1.

#### Assessment of gene copy numbers by quantitative real-time PCR

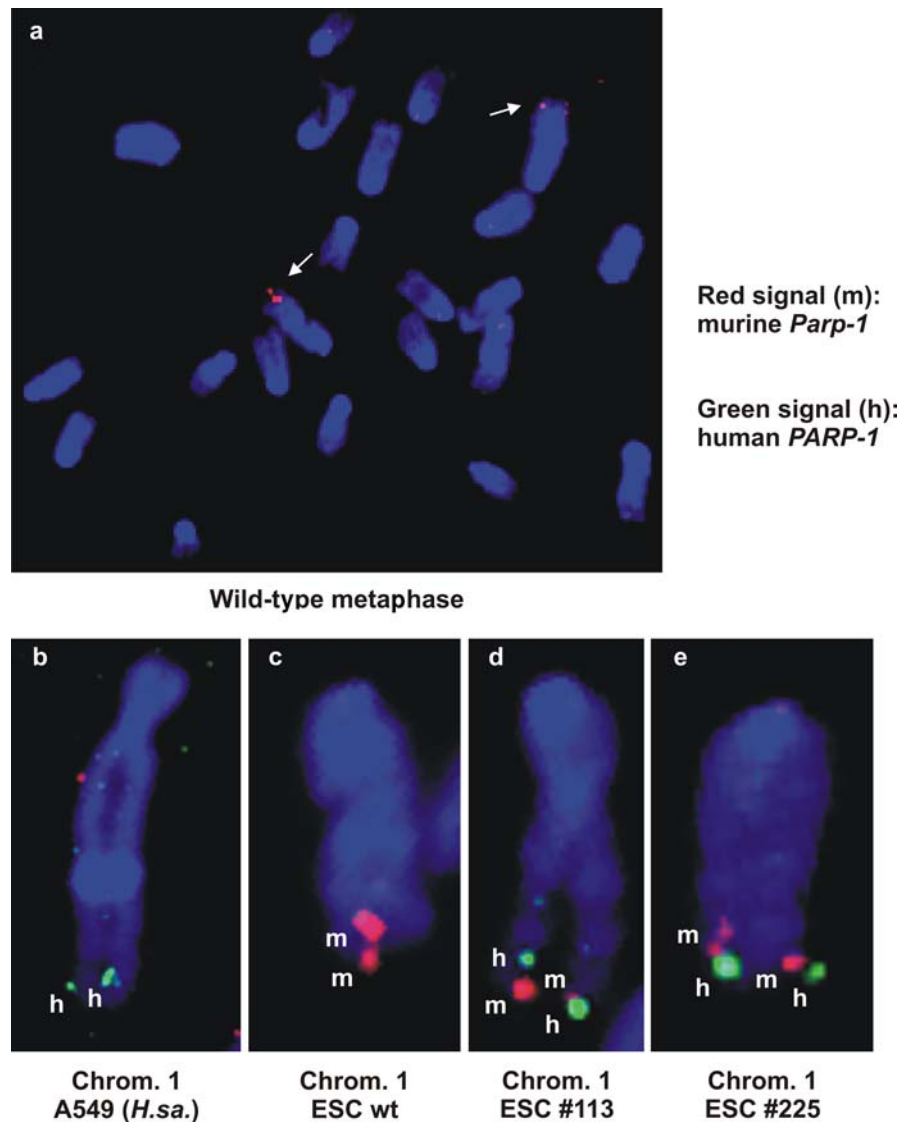
Quantitative real-time PCR (qPCR) analyses of genomic DNA were performed: first, to verify the results from FISH analyses; second, to identify potential *hPARP-1* homozygote animals; and third, to analyze if potential extension of the vector homology arms, which would have been another possible explanation for the results obtained by flanking PCRs and Southern blot analyses, affected gene copy numbers of *mParp-1*-flanking genes.

Gene copy numbers of *mParp-1*, *hPARP-1*, *Gm821* (GenBank ID: 329305) (proximal of *mParp-1*), and *Lin9* (GenBank ID: 72568) (distal of *mParp-1*) were analyzed by qPCR. Validation of qPCRs is shown in Fig. S5a.

For the analysis of *hPARP-1* gene copy numbers, the *hPARP-1* gene copy number of the diploid human fibroblast cell line IMR90 was set to 2.0. Wild-type E14-1 ES cells and C57BL/6 mice showed equal *mParp-1* gene copy numbers, thereby ruling out a duplication of the *mParp-1* locus in the wild-type E14-1 ES cells a priori. Therefore, murine wild-type genomic DNA was used as a reference with *mParp-1* gene copy numbers set to 2.0.

Consistent with the results from FISH analyses, but prima facie contradictory to the results from flanking PCR and Southern blot analyses, the presence of one copy of *hPARP-1* and two gene copies of *mParp-1* were detected in the genome of ES cell clones #113, #225, and #267 (Fig. 4a, b). As expected, fully heterozygous *hPARP-1* mice of the F1 generation showed one gene copy of *hPARP-1*. However, in the F2 generation, which was obtained by intercrossing of *hPARP-1* heterozygous mice, *hPARP-1* homozygotes with two copies of *hPARP-1* were identified (Fig. 4a). These findings were confirmed by analyzing signal intensities of Southern blots (see Fig. 5b) and by breeding experiments: Mice from lines #113 and #225 ( $n = 5$ ) that were identified as *hPARP-1* homozygotes by qPCR produced litters comprising 100% mutant animals after breeding to wild-type animals as expected in case of uniparental homozygosity (see Fig. S5b).

**Fig. 3** Fluorescence in situ hybridization (FISH) of targeted ES cells. Hybridization probes specific for the *hPARP-1* (green signal; *h*) or *mParp-1* (red signal; *m*) were amplified by PCR methods and labeled with biotin-dUTP or digoxigenin-dUTP by nick translation, respectively. DNA was visualized by DAPI staining. **a** Overview of a metaphase spread demonstrating the specificity of the hybridization probes to chromosome 1. **b** and **c** On chromosomes 1 from human A549 cells or wild-type murine ES cells only signals from the *hPARP-1* or from the *mParp-1* probes were detectable, respectively. **d** and **e** Hybridization signals on chromosomes 1 of mutant murine ES cell clones #113 and #225, respectively. ESC, ES cell clone; *H.sa.*, *Homo sapiens*

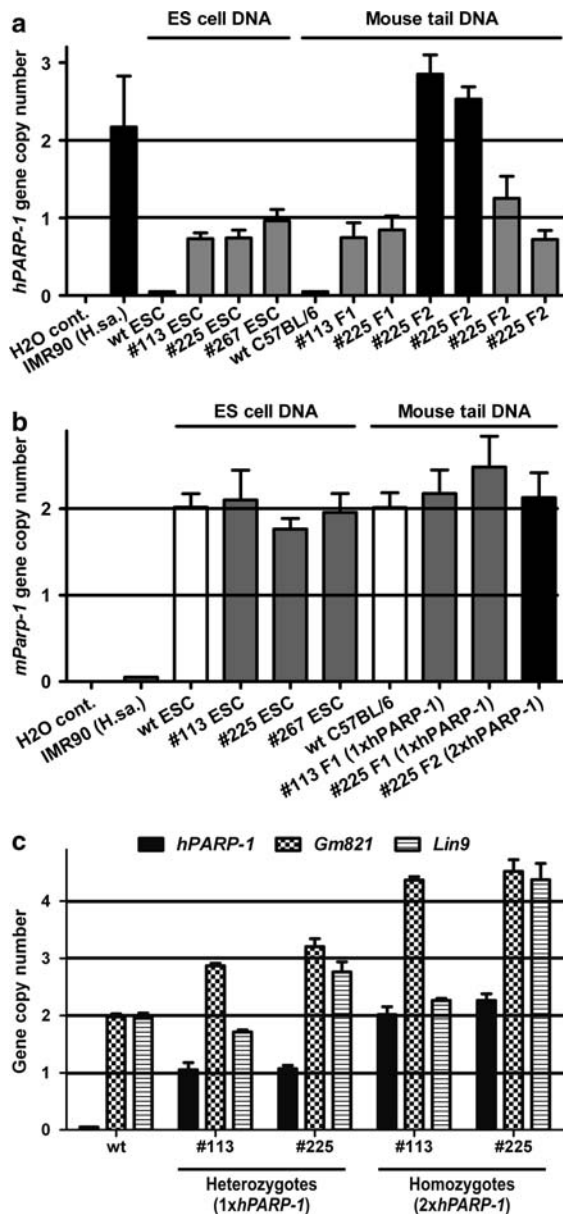


Determination of gene copy numbers of 4 litters of line #113 and 6 litters of line #225 revealed a distribution of genotypes (wt,  $1 \times hPARP-1$ ,  $2 \times hPARP-1$ ) consistent with a Mendelian distribution (Table 1b).

Taken together, data obtained from animal breeding, FISH, and qPCR strongly indicated that in all three independently generated ES cell clones an ectopic gene targeting event had taken place. Such an event can be explained as a gene conversion by the SDSA pathway of DSB repair, in which the targeting vector homology arms invaded the endogenous *mParp-1* locus at the promoter and the terminator sites. Murine *Parp-1* sequences were then copied

onto their ends, and the vector subsequently integrated at an adjacent non-homologous position. Due to copying of flanking non-homologous sequences this is registered as a bona fide site-specific homologous recombination. As a result, homozygous *hPARP-1* mice carry two copies of the endogenous *mParp-1* locus plus two ectopic copies of the *hPARP-1* coding sequence.

In order to investigate to what extent flanking regions were extended and if other loci could be affected, we analyzed possible gene duplications of the *mParp-1*-flanking genes *Gm821* (50 kb proximal of the *mParp-1* locus) and *Lin9* (40 kb distal of the *mParp-1* locus) by qPCR (Fig. 4c). Our analyses



**Fig. 4** Determination of gene copy numbers in mutant ES cells and mice by quantitative real-time PCR (qPCR). **a** Human *PARP-1* gene copy numbers as determined from genomic DNA of ES cell clones and mouse tail biopsies. Human diploid fibroblasts (IMR90) were used as standard with *hPARP-1* gene copy number set to 2.0. **b** Murine *Parp-1* gene copy numbers as determined from genomic DNA of ES cell clones and mouse tail biopsies. Murine wild-type ES cells were used as a standard and *mParp-1* gene copy number was set to 2.0. **c** Determination of gene copy numbers of *mParp-1*-flanking genes. Genomic DNA from six wild-type, four heterozygous, and three homozygous *hPARP-1* mice of mouse lines #113 and #225, respectively, were analyzed for *hPARP-1*, *Gm821*, and *Lin9* gene copy numbers. Analysis of each mouse was performed in technical triplicates. *Gm821* and *Lin9* flank the *mParp-1* locus on its proximal or distal side, respectively. In the case of *Gm821* and *Lin9* gene copy numbers of wild-type mice were set to 2.0. Means  $\pm$  SEM. #113, #225, #267 indicate mutant ES cell clones or mice. Mice of F2 were obtained by interbreeding of *hPARP-1* heterozygous mice of F1. 1 $\times$ *hPARP-1*, heterozygous mutant mice; 2 $\times$ *hPARP-1*, homozygous mutant mice; cont., control; ESC, ES cell clone; F1 and F2, first and second filial generation; H.sa., *Homo sapiens*

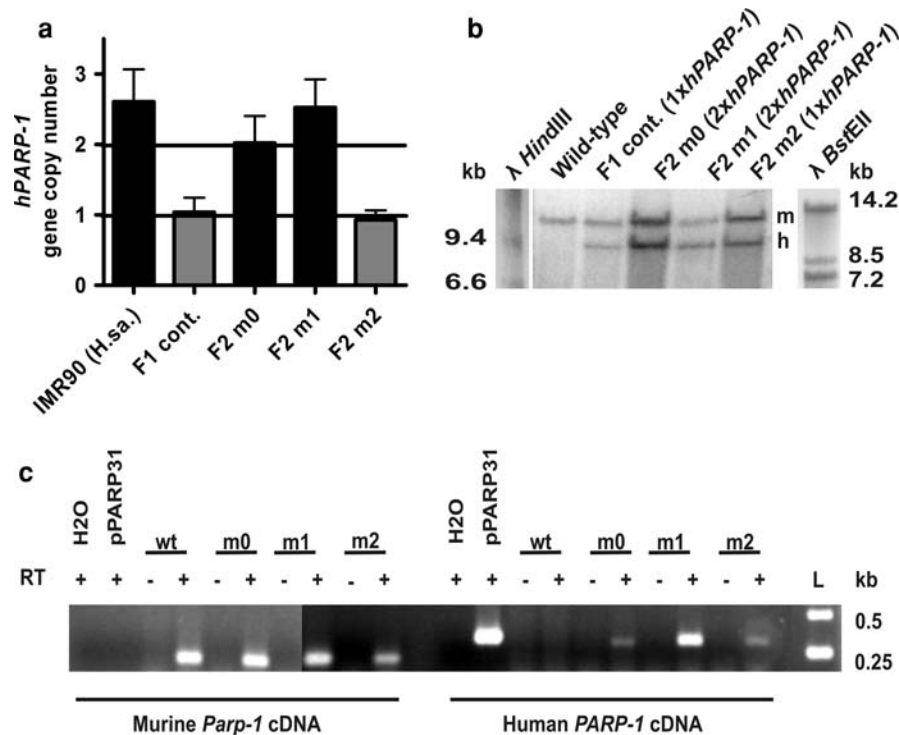
elongated to more than 40 kb by copying of non-homologous *mParp-1*-flanking sequences.

#### Expression analyses

To test if the integrated *hPARP-1* coding sequence and the remaining *mParp-1* gene are transcriptionally functional, we performed mRNA analyses by RT-PCR with primer pairs specific either for the murine or the human transcript. The results demonstrated the presence of both transcripts in heterozygous as well as homozygous *hPARP-1* mice (Fig. 5). This suggested that the *mParp-1* locus on the targeted chromosome was functional and not destroyed during gene targeting.

To investigate if expression of mRNA gave rise to the respective full-length proteins, we examined total, human, and murine PARP-1 expression, respectively, by western blot analyses of spleen extracts. Our results demonstrated a gene dose-dependent, increased of total PARP-1 expression in both mouse lines generated (#113 and #225) (Fig. 6). Furthermore, both mutant lines displayed expression of full-length hPARP-1 in a gene dose-dependent manner. Analysis for mPARP-1, as assessed by using a novel mPARP-1-specific polyclonal antibody, showed full-length murine PARP-1 expression confirming expression analyses on the mRNA level. Additional bands of truncated

revealed that *hPARP-1* mice of lines #113 and #225 carry duplicated regions of *Gm821*, which corresponds to three gene copies in heterozygotes and four gene copies in homozygotes. In addition, line #225 carries a duplication of *Lin9*. The difference in the extension length of the terminator homology arm of lines #113 and #225 verifies that the two lines were generated in two independent gene targeting events. Furthermore, these results demonstrate that in line #225 the vector homology arms were bidirectionally



**Fig. 5** PARP-1 transcript expression analysis in mutant mice. **a** Human *PARP-1* gene copy numbers of three *hPARP-1* mice of line #225 F2 (m0, m1, m2) used for this experiment as determined by qPCR. Diploid human fibroblasts (*IMR90*) were used as standard with gene copy number set to 2.0. Means  $\pm$  SEM,  $n = 3$ . **b** Southern blot analysis of genomic DNA from the same mice as shown in **a**. Kidney-DNA was digested with *BsrGI* and terminator fragments detected by hybridization with an ‘external’  $^{32}\text{P}$ -labeled probe (probe d, see Fig. 3a). Note the different ratios of signal intensities of *mParp-1* to *hPARP-1* in mutant heterozygous ( $\sim 2:1$ ) and

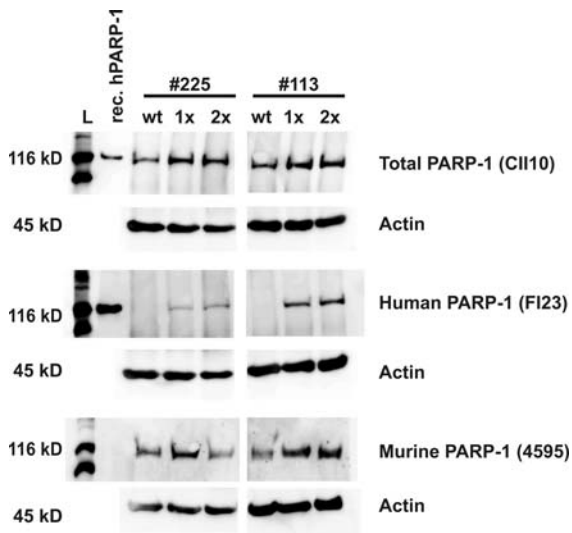
homozygous mice ( $\sim 1:1$ ). *h* indicates molecular size of the *hPARP-1* knock-in fragments; *m* of the *mParp-1* wild-type fragments. **c** Qualitative RT-PCR from mRNA isolated from spleens of mice m0, m1, and m2. Specific fragments from corresponding cDNA were amplified using primer pairs specific for *mParp-1* (left) or *hPARP-1* (right), respectively. Plasmid DNA pPARP31, carrying the *hPARP-1* cDNA was used as specificity control. *F1* and *F2*, first and second filial generation; *H.sa.*, *Homo sapiens*; *L*, molecular size ladder; *m*, male; *RT*, reverse transcriptase

proteins were not detected indicating that the endogenous *mParp-1* and the ectopic *hPARP-1* loci are fully functional.

## Discussion

In order to generate human *PARP-1* knock-in mice with complete replacement of the murine *Parp-1* coding sequence (32 kb) with its human orthologous sequence (46 kb), we performed gene targeting in ES cells using an ‘ends-out’ gene replacement vector (van der Weyden et al. 2002) (Fig. S1). The *hPARP-1* knock-in sequence was flanked by two homology regions comprising isogenic sequences of 5 kb each from the *mParp-1* promoter and terminator region,

which permit site-specific homologous recombination. Although we are not aware of a report showing gene replacement of that size (32 kb vs. 46 kb) by traditional gene targeting, we favored this method over that of recombination-mediated cassette exchange (RMCE), which utilizes the Cre-*loxP* system, because in this case at least two successive rounds of gene targeting would have been necessary (Yu and Bradley 2001; Roebroek et al. 2003; Brault et al. 2006; Wallace et al. 2007). Moreover, our approach was theoretically feasible, since the homology arms possessed a reasonable quality, i.e., sequences were isogenic to those of the E14-1 ES cell line, and size, i.e., 5 kb (Hasty et al. 1991; Deng and Capecchi 1992), and it was shown that gene replacement vectors are able to generate deletions of genomic sequences of up to 30 kb (Zhang et al.



**Fig. 6** PARP-1 expression in mutant mice. Western blot analyses for PARP-1 expression using spleen lysates prepared from wild-type, mutant heterozygous (“1x”) and homozygous (“2x”) animals of lines #225 and #113 as identified by qPCR. The upper blot was probed with general-PARP-1-specific antibody CII10, the blot in the middle with human-PARP-1-specific antibody FI23, and the blot at the bottom with murine-PARP-1-specific antibody 4595. Detection of actin was used as a loading control. *L*, molecular weight ladder; *rec.*, recombinant

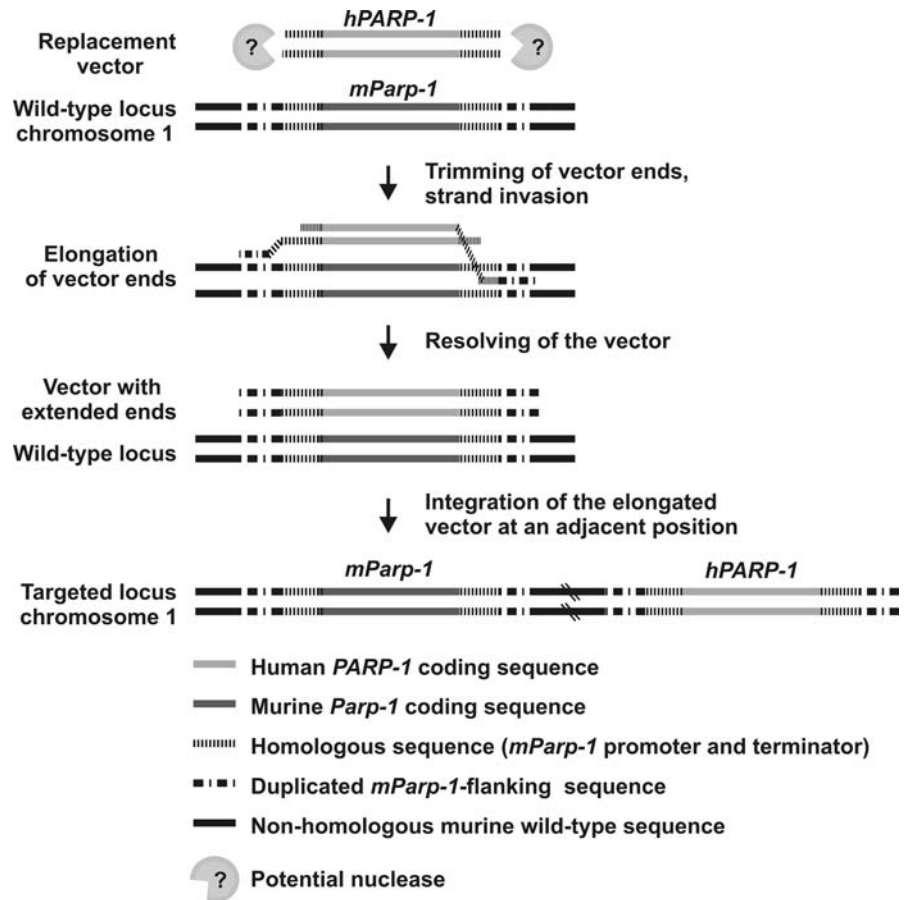
1994; Brault et al. 2006). Furthermore, in previous studies a BAC-derived replacement-type targeting construct was used to introduce two mutations in the mouse genome as far as 43 kb apart in a single round of gene targeting (Testa et al. 2003). After positive/negative selection of electroporated ES cells, we analyzed at total of 276 G418-resistant clones for site-specific HR by flanking PCRs of the promoter and terminator homology arms and sequencing of the obtained amplicons (Figs. 1, S3). Subsequently, we verified these results by Southern blotting with internal and external probes (Figs. 2, 5b). Thereby, we identified three clones (1.1%) with bona fide site-specific recombination of the targeting vector. This targeting efficacy was in the lower range of that one would expect with a replacement vector of this size (Zhang et al. 1994; Testa et al. 2003). Despite the fact that we had obtained bona fide site-specific HR in three ES cell clones picked from three different G418-selected culture plates, subsequent mouse breeding (Table 1), FISH analyses (Fig. 3) and qPCR (Fig. 4b) unexpectedly revealed a remaining copy of the endogenous *mParp-1* in close proximity to an ectopic single copy of

*hPARP-1* on chromosome 1. In addition, expression analyses both on the mRNA and protein level demonstrated full-length murine and human PARP-1 expression in mutant homozygotes in two out of two analyzed lines (#113 and #225) (Figs. 5, 6).

The theoretical possibility that the *mParp-1* locus was duplicated a priori in the wild-type E14-1 ES cells can be ruled out, because the *mParp-1* gene copy number of E14-1 ES cells equals that of wild-type C57BL/6 mice and there was no decrease in *mParp-1* gene copy numbers observed upon gene targeting in ES cell clones #113, #225, and #267 (Fig. 4b). In accordance with this finding, FISH analyses excluded the existence of an additional chromosome 1 (trisomy 1). Initially, we hypothesized that *mParp-1* pseudogenes might have been targeted. Interestingly, *PARP-1* pseudogenes are present in the human genome on chromosomes 13 and 14 (Cherney et al. 1987; Baumgartner et al. 1992). However, our BLAST searches did not yield any evidence for the presence of a *mParp-1* pseudogene or extended homologies in the mouse genome. Another possible explanation is based on the fact that large DNA fragments are prone to breakage by hydrodynamic shearing (Yang and Seed 2003). But HR of multiple fragments of the targeting construct used in this study can be excluded, because we detected only full-length PARP-1 proteins, of human as well as murine origin, in our ES clones.

A further possibility to interpret our results refers to an observation so far only described with ‘ends-in’ insertion-type vectors in non-ES-cell mammalian gene targeting (Adair et al. 1989, 1990, 1998; Itzhaki and Porter 1991; Scheerer and Adair 1994; Dellaire et al. 1997; McCulloch et al. 2003) and designated ectopic gene targeting, or ‘pick-up’ event (Dorin et al. 1992; Dellaire et al. 1997). Based on our data and in line with reports of ectopic gene targeting, we propose the following model to explain the outcome of the gene targeting approach observed in this study: upon electroporation, the homologous regions of the arms of the targeting vector were made single-stranded by 5’–3’ nuclease activity and the resulting single-stranded 3’ ends invaded the endogenous wild-type locus (Fig. 7; Sun et al. 1991; Barzel and Kupiec 2008). DNA polymerases then extended the arms bidirectionally for at least 2 kb using endogenous sequences as templates, as shown by DNA sequencing of flanking PCR amplicons (Figs. 1, S3) and Southern

**Fig. 7** Proposed model explaining the outcome of the gene targeting experiment described. The indicated integration position of the targeting vector is speculative. Illustration not drawn to scale. For details, see text



blot analyses (Figs. 2, 5b). Quantitative PCR of *mParp-1*-flanking regions (*Gm821* and *Lin9*) indicated that extension occurred to >50 kb onto the promoter homology arm of clone #113 and to >40 kb onto both homology arms, i.e., promoter and terminator, of clone #225 (Fig. 4c). Subsequently, the whole vector integrated at an adjacent position. Strikingly, the mechanism of elongation of the homology arm is consistent with the SDSA pathway, which was described as the predominant mechanism of cellular DSB repair (Helleday et al. 2007).

The phenomenon of ectopic gene targeting was first described in Chinese hamster ovary (CHO) cells in the context of targeting of the adenine phosphoribosyl transferase (*APRT*) locus and after retroviral transduction of rat cells (Adair et al. 1989; Ellis and Bernstein 1989). Several reports followed, demonstrating ectopic gene targeting in CHO, murine fibroblasts, and hybridoma cells, respectively, but so far not in murine ES cells (Adair et al. 1990, 1998; Itzhaki and Porter 1991; Aratani et al. 1992; Scheerer

and Adair 1994). All these experiments had in common that ‘ends-in’ insertion-type vectors were used and elongation of the targeting vector occurred in a unidirectional manner from a single resected 3’ end of the vector (Belmaaza and Chartrand 1994). The first report of a bidirectional elongation of an ‘ends-in’ integration vector in mammalian cells was published by McCulloch and colleagues in a mammalian hybridoma cell line (McCulloch et al. 2003). They also showed that extension of the vector ends can exceed 16 kb in length. Interestingly, ectopic gene targeting with unidirectional extension of one homology arm was shown to occur at a frequency similar to that observed in traditional gene targeting (Aratani et al. 1992; Dellaire et al. 1997). Therefore, it is rather surprising that in the current literature no such an event was described, when using ‘ends-out’ gene replacement vectors in traditional gene targeting in murine ES cells (McCulloch et al. 2003). McCulloch and colleagues suggested that the free DNA ends in the ‘ends-out’ configuration may behave differently from those

in the ‘ends-in’ configuration as a consequence of potential differences in the way the two types of ends may be regulated during recombination (McCulloch et al. 2003). However, in the present study, we report, to our knowledge, the first ectopic gene targeting event with bidirectional copying of endogenous sequences exceeding 40 kb onto the ends of an ‘ends-out’ gene replacement vector in murine ES cells. This study therefore provides a proof-of-principle by showing that in addition to Holliday junction formation also the SDSA pathway of the cellular DSB repair machinery can be involved in the processes that occur during mammalian gene targeting. The reason, why in the present study ectopic gene targeting occurred in three out of three identified ES cell clones and therefore represented the rule rather than an exception remains elusive. One possible explanation could be given by the size of the targeting vector of 64 kb. Targeting vectors of this size are a recent development and are not studied as extensively as compared to smaller targeting vectors of up to 20 kb. Moreover, the nature of the targeted locus itself could play a role, as it might favor SDSA over the resolution of the Holliday junction by two crossing-over events as is the case in traditional gene targeting. In this regard, it could be speculated that PARP-1 itself may play a role in such a scenario. Thus, it was shown that PARP-1 can bind to promoter sequences including its own promoter and act as a modulator of transcriptional promoter activity (Oei et al. 1994; Soldatenkov et al. 2002; Lonskaya et al. 2005; Potaman et al. 2005). Moreover, the role of PARP-1 in DNA repair and genomic stability is well established (Schreiber et al. 2006). It has been shown that HR leading to sister chromatid exchange, is either inhibited by overexpression of PARP-1 or increased by its pharmacological inhibition (Sonoda et al. 1999; Dominguez et al. 2000; Meyer et al. 2000). Since PARP-1 has been found to exert anti-recombinogenic activity by supporting a switch to the non-homologous end joining pathway of DSB repair, it could be speculated that the presence of PARP-1 at its own gene locus might have counteracted gene replacement (Waldman and Waldman 1990, 1991; Semionov et al. 2003; Dominguez-Bendala et al. 2006). However, one should keep in mind that PARP-1 is a very abundant nuclear protein (D’Amours et al. 1999) especially in ES cells (unpublished data) and that its antirecombinogenic effect could also affect other gene

loci (Semionov et al. 2003; Dominguez-Bendala et al. 2006).

Dellaire and colleagues showed that unidirectional ectopic gene targeting in murine fibroblasts followed a bimodal distribution. They demonstrated that the targeting vector either integrated non-randomly within 3 Mb of the target gene, or on another chromosome in a random fashion, as they speculated (Dellaire et al. 1997). In the present study, the question of random versus non-random ectopic integration of the elongated targeting vector could so far not fully be clarified, because FISH analyses revealed that the ectopic *hPARP-1* copy integrated in close proximity to endogenous *mParp-1* locus on chromosome 1 without showing a clear tendency for a distal or proximal location (Fig. 3). Since the technical resolution of FISH analysis is limited to 2–3 Mb (Trask 1991), this indicates that the *hPARP-1*-targeted site lies within kb up to several Mb from the endogenous *mParp-1* locus. Although, it was shown that targeting efficacy increases with the length of the homology arms, as little as 14–25 bp of homology are sufficient to trigger integration of the construct by HR (Rubnitz and Subramani 1984; Ayares et al. 1986; Hasty et al. 1991). If in the three identified targeted ES cell clones the elongated vector integrated randomly or in a non-random fashion at a site of a putative microhomology or a genetic locus that might be especially prone to integration of the extrachromosomal DNA due to its chromatin structure, remains a matter of speculation.

## Conclusions

In summary, in this study, we present the generation of a novel mouse model with expression of the human PARP-1 by ectopic gene targeting in murine ES cells using an ‘ends-out’ gene replacement vector. Unexpectedly, bidirectional elongation of the vector homology arms, in some cases to more than 40 kb, and subsequent integration of the targeting vector at an adjacent position on chromosome 1 mimicked site-specific HR and gene replacement as determined by flanking PCRs and Southern blot analyses. To our knowledge, this report demonstrates for the first time that, as predicted by the DSB repair model, the SDSA pathway can also operate during traditional gene

targeting in ES cells. This finding could be of general importance, as it potentially represents a critical caveat for future gene targeting approaches, especially when generating knock-in mice. Typically, in this context loss of protein expression of the endogenous targeted gene cannot easily be assessed during ES cell screening. Here we present two reliable techniques, i.e., qPCR and FISH analysis, to detect ectopic gene targeting *in vitro* before attempting to generate homozygous knock-in mice by ES cell injection into blastocysts.

**Acknowledgments** We thank Oliver Popp, Gudrun von Scheven, Daniela Gassen, Heidi Henseleit, and Birgitt Planitz for technical help and support, and Prof. Andrew Smith (Edinburgh, UK) for valuable discussion of the data. AM was supported by the ‘Studienstiftung des Deutschen Volkes’ and the ‘Deutsche Forschungs Gemeinschaft’ (DFG) (International Research Training Research Group 1331).

## References

- Adair GM, Nairn RS, Wilson JH et al (1989) Targeted homologous recombination at the endogenous adenine phosphoribosyltransferase locus in Chinese hamster cells. *Proc Natl Acad Sci USA* 86:4574–4578. doi:[10.1073/pnas.86.12.4574](https://doi.org/10.1073/pnas.86.12.4574)
- Adair GM, Nairn RS, Wilson JH et al (1990) Targeted gene replacement at the endogenous APRT locus in CHO cells. *Somat Cell Mol Genet* 16:437–441. doi:[10.1007/BF01233193](https://doi.org/10.1007/BF01233193)
- Adair GM, Scheerer JB, Brotherman A et al (1998) Targeted recombination at the Chinese hamster APRT locus using insertion versus replacement vectors. *Somat Cell Mol Genet* 24:91–105. doi:[10.1023/B:SCAM.0000007112.62928.d8](https://doi.org/10.1023/B:SCAM.0000007112.62928.d8)
- Aratani Y, Okazaki R, Koyama H (1992) End extension repair of introduced targeting vectors mediated by homologous recombination in mammalian cells. *Nucleic Acids Res* 20:4795–4801. doi:[10.1093/nar/20.18.4795](https://doi.org/10.1093/nar/20.18.4795)
- Ayares D, Chekuri L, Song KY et al (1986) Sequence homology requirements for intermolecular recombination in mammalian cells. *Proc Natl Acad Sci USA* 83:5199–5203. doi:[10.1073/pnas.83.14.5199](https://doi.org/10.1073/pnas.83.14.5199)
- Babinet C, Cohen-Tannoudji M (2001) Genome engineering via homologous recombination in mouse embryonic stem (ES) cells: an amazingly versatile tool for the study of mammalian biology. *An Acad Bras Cienc* 73:365–383
- Barzel A, Kupiec M (2008) Finding a match: how do homologous sequences get together for recombination? *Nat Rev Genet* 9:27–37. doi:[10.1038/nrg2224](https://doi.org/10.1038/nrg2224)
- Baumgartner M, Schneider R, Auer B et al (1992) Fluorescence *in situ* mapping of the human nuclear NAD<sup>+</sup> ADP-ribosyltransferase gene (ADPRT) and two secondary sites to human chromosomal bands 1q42, 13q34, and 14q24. *Cytogenet Cell Genet* 61:172–174. doi:[10.1159/000133400](https://doi.org/10.1159/000133400)
- Belmaaza A, Chartrand P (1994) One-sided invasion events in homologous recombination at double-strand breaks. *Mutat Res* 314:199–208
- Braut V, Pereira P, Duchon A et al (2006) Modeling chromosomes in mouse to explore the function of genes, genomic disorders, and chromosomal organization. *PLoS Genet* 2:e86. doi:[10.1371/journal.pgen.0020086](https://doi.org/10.1371/journal.pgen.0020086)
- Capecchi MR (2005) Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nat Rev Genet* 6:507–512. doi:[10.1038/nrg1619](https://doi.org/10.1038/nrg1619)
- Cherney BW, McBride OW, Chen DF et al (1987) cDNA sequence, protein structure, and chromosomal location of the human gene for poly(ADP-ribose) polymerase. *Proc Natl Acad Sci USA* 84:8370–8374. doi:[10.1073/pnas.84.23.8370](https://doi.org/10.1073/pnas.84.23.8370)
- Collins FS, Rossant J, Wurst W (2007) A mouse for all reasons. *Cell* 128:9–13. doi:[10.1016/j.cell.2006.12.018](https://doi.org/10.1016/j.cell.2006.12.018)
- Copeland NG, Jenkins NA, Court DL (2001) Recombinering: a powerful new tool for mouse functional genomics. *Nat Rev Genet* 2:769–779. doi:[10.1038/35093556](https://doi.org/10.1038/35093556)
- D’Amours D, Desnoyers S, D’Silva I et al (1999) Poly(ADP-ribosylation) reactions in the regulation of nuclear functions. *Biochem J* 342(Pt 2):249–268. doi:[10.1042/0264-6021:3420249](https://doi.org/10.1042/0264-6021:3420249)
- Dellaire G, Lemieux N, Belmaaza A et al (1997) Ectopic gene targeting exhibits a bimodal distribution of integration in murine cells, indicating that both intra- and interchromosomal sites are accessible to the targeting vector. *Mol Cell Biol* 17:5571–5580
- Deng C, Capecchi MR (1992) Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. *Mol Cell Biol* 12:3365–3371
- Dominguez I, Mateos S, Cortes F (2000) Yield of SCEs and translocations produced by 3-aminobenzamide in cultured Chinese hamster cells. *Mutat Res* 448:29–34. doi:[10.1016/S0027-5107\(99\)00228-6](https://doi.org/10.1016/S0027-5107(99)00228-6)
- Dominguez-Bendala J, Masutani M, McWhir J (2006) Down-regulation of PARP-1, but not of Ku80 or DNA-PKcs’, results in higher gene targeting efficiency. *Cell Biol Int* 30:389–393. doi:[10.1016/j.cellbi.2005.12.005](https://doi.org/10.1016/j.cellbi.2005.12.005)
- Dorin JR, Emslie E, Hanratty D et al (1992) Gene targeting for somatic cell manipulation: rapid analysis of reduced chromosome hybrids by Alu-PCR fingerprinting and chromosome painting. *Hum Mol Genet* 1:53–59. doi:[10.1093/hmg/1.1.53](https://doi.org/10.1093/hmg/1.1.53)
- Elefanty AG, Begley CG, Metcalf D et al (1998) Characterization of hematopoietic progenitor cells that express the transcription factor SCL, using a lacZ “knock-in” strategy. *Proc Natl Acad Sci USA* 95:11897–11902. doi:[10.1073/pnas.95.20.11897](https://doi.org/10.1073/pnas.95.20.11897)
- Ellis J, Bernstein A (1989) Gene targeting with retroviral vectors: recombination by gene conversion into regions of nonhomology. *Mol Cell Biol* 9:1621–1627
- Glaser S, Anastassiadis K, Stewart AF (2005) Current issues in mouse genome engineering. *Nat Genet* 37:1187–1193. doi:[10.1038/ng1668](https://doi.org/10.1038/ng1668)
- Hasty P, Rivera-Perez J, Bradley A (1991) The length of homology required for gene targeting in embryonic stem cells. *Mol Cell Biol* 11:5586–5591

- Helleday T, Lo J, van Gent DC et al (2007) DNA double-strand break repair: from mechanistic understanding to cancer treatment. *DNA Repair (Amsterdam)* 6:923–935. doi: [10.1016/j.dnarep.2007.02.006](https://doi.org/10.1016/j.dnarep.2007.02.006)
- Itzhaki JE, Porter AC (1991) Targeted disruption of a human interferon-inducible gene detected by secretion of human growth hormone. *Nucleic Acids Res* 19:3835–3842. doi: [10.1093/nar/19.14.3835](https://doi.org/10.1093/nar/19.14.3835)
- Lee EC, Yu D, Martinez de Velasco J et al (2001) A highly efficient *Escherichia coli*-based chromosome engineering system adapted for recombinogenic targeting and sub-cloning of BAC DNA. *Genomics* 73:56–65. doi: [10.1006/geno.2000.6451](https://doi.org/10.1006/geno.2000.6451)
- Li J, Baker MD (2000) Mechanisms involved in targeted gene replacement in mammalian cells. *Genetics* 156:809–821
- Li X, Heyer WD (2008) Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res* 18:99–113. doi: [10.1038/cr.2008.1](https://doi.org/10.1038/cr.2008.1)
- Li J, Read LR, Baker MD (2001) The mechanism of mammalian gene replacement is consistent with the formation of long regions of heteroduplex DNA associated with two crossing-over events. *Mol Cell Biol* 21:501–510. doi: [10.1128/MCB.21.2.501-510.2001](https://doi.org/10.1128/MCB.21.2.501-510.2001)
- Lonskaya I, Potaman VN, Shlyakhtenko LS et al (2005) Regulation of poly(ADP-ribose) polymerase-1 by DNA structure-specific binding. *J Biol Chem* 280:17076–17083. doi: [10.1074/jbc.M413483200](https://doi.org/10.1074/jbc.M413483200)
- Luche H, Weber O, Nageswara Rao T et al (2007) Faithful activation of an extra-bright red fluorescent protein in “knock-in” Cre-reporter mice ideally suited for lineage tracing studies. *Eur J Immunol* 37:43–53. doi: [10.1002/eji.200636745](https://doi.org/10.1002/eji.200636745)
- Luo JL, Yang Q, Tong WM et al (2001) Knock-in mice with a chimeric human/murine p53 gene develop normally and show wild-type p53 responses to DNA damaging agents: a new biomedical research tool. *Oncogene* 20:320–328. doi: [10.1038/sj.onc.1204080](https://doi.org/10.1038/sj.onc.1204080)
- Mak TW (2007) Gene targeting in embryonic stem cells scores a knockout in Stockholm. *Cell* 131:1027–1031. doi: [10.1016/j.cell.2007.11.033](https://doi.org/10.1016/j.cell.2007.11.033)
- McCulloch RD, Read LR, Baker MD (2003) Strand invasion and DNA synthesis from the two 3' ends of a double-strand break in mammalian cells. *Genetics* 163:1439–1447
- Meyer R, Muller M, Beneke S et al (2000) Negative regulation of alkylation-induced sister-chromatid exchange by poly(ADP-ribose) polymerase-1 activity. *Int J Cancer* 88: 351–355. doi: [10.1002/1097-0215\(20001101\)88:3<351::AID-IJC5>3.0.CO;2-H](https://doi.org/10.1002/1097-0215(20001101)88:3<351::AID-IJC5>3.0.CO;2-H)
- Nagy A, Gertsenstein M, Vintersten K, Behringer R (2003) *Manipulating the mouse embryo*, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- Oei SL, Herzog H, Hirsch-Kauffmann M et al (1994) Transcriptional regulation and autoregulation of the human gene for ADP-ribosyltransferase. *Mol Cell Biochem* 138:99–104. doi: [10.1007/BF00928449](https://doi.org/10.1007/BF00928449)
- Potaman VN, Shlyakhtenko LS, Oussatcheva EA et al (2005) Specific binding of poly(ADP-ribose) polymerase-1 to cruciform hairpins. *J Mol Biol* 348:609–615. doi: [10.1016/j.jmb.2005.03.010](https://doi.org/10.1016/j.jmb.2005.03.010)
- Roebroek AJ, Wu X, Bram RJ (2003) Knockin approaches. *Methods Mol Biol* 209:187–200
- Rozmahel R, Gyomory K, Plyte S et al (1997) Incomplete rescue of cystic fibrosis transmembrane conductance regulator deficient mice by the human CFTR cDNA. *Hum Mol Genet* 6:1153–1162. doi: [10.1093/hmg/6.7.1153](https://doi.org/10.1093/hmg/6.7.1153)
- Rubnitz J, Subramani S (1984) The minimum amount of homology required for homologous recombination in mammalian cells. *Mol Cell Biol* 4:2253–2258
- Sambrook A (2001) *Molecular cloning: a laboratory manual*, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- Scheerer JB, Adair GM (1994) Homology dependence of targeted recombination at the Chinese hamster APRT locus. *Mol Cell Biol* 14:6663–6673
- Scherthan H, Seisenberger C, Greulich K et al (1994) Mapping of the murine nuclear factor I/X gene (Nfix) to mouse chromosome 8 C1-2 by FISH. *Genomics* 22:247–249
- Schreiber V, Dantzer F, Ame J-C et al (2006) Poly(ADP-ribose): novel functions for an old molecule. *Nat Rev Mol Cell Biol* 7:517–528. doi: [10.1038/nrm1963](https://doi.org/10.1038/nrm1963)
- Semionov A, Cournoyer D, Chow TY (2003) 1,5-Isoquinolediol increases the frequency of gene targeting by homologous recombination in mouse fibroblasts. *Biochem Cell Biol* 81:17–24. doi: [10.1139/o02-172](https://doi.org/10.1139/o02-172)
- Soldatenkov VA, Chasovskikh S, Potaman VN et al (2002) Transcriptional repression by binding of poly(ADP-ribose) polymerase to promoter sequences. *J Biol Chem* 277:665–670. doi: [10.1074/jbc.M108551200](https://doi.org/10.1074/jbc.M108551200)
- Sonoda E, Sasaki MS, Morrison C et al (1999) Sister chromatid exchanges are mediated by homologous recombination in vertebrate cells. *Mol Cell Biol* 19:5166–5169
- Sun H, Treco D, Szostak JW (1991) Extensive 3'-overhanging, single-stranded DNA associated with the meiosis-specific double-strand breaks at the ARG4 recombination initiation site. *Cell* 64:1155–1161. doi: [10.1016/0092-8674\(91\)90270-9](https://doi.org/10.1016/0092-8674(91)90270-9)
- Testa G, Zhang Y, Vintersten K et al (2003) Engineering the mouse genome with bacterial artificial chromosomes to create multipurpose alleles. *Nat Biotechnol* 21:443–447. doi: [10.1038/nbt804](https://doi.org/10.1038/nbt804)
- Trask BJ (1991) Fluorescence in situ hybridization: applications in cytogenetics and gene mapping. *Trends Genet* 7:149–154
- van der Weyden L, Adams DJ, Bradley A (2002) Tools for targeted manipulation of the mouse genome. *Physiol Genomics* 11:133–164
- Waldman BC, Waldman AS (1990) Illegitimate and homologous recombination in mammalian cells: differential sensitivity to an inhibitor of poly(ADP-ribosylation). *Nucleic Acids Res* 18:5981–5988. doi: [10.1093/nar/18.20.5981](https://doi.org/10.1093/nar/18.20.5981)
- Waldman AS, Waldman BC (1991) Stimulation of intrachromosomal homologous recombination in mammalian cells by an inhibitor of poly(ADP-ribosylation). *Nucleic Acids Res* 19:5943–5947. doi: [10.1093/nar/19.21.5943](https://doi.org/10.1093/nar/19.21.5943)
- Wallace HA, Marques-Kranc F, Richardson M et al (2007) Manipulating the mouse genome to engineer precise functional syntenic replacements with human sequence. *Cell* 128:197–209. doi: [10.1016/j.cell.2006.11.044](https://doi.org/10.1016/j.cell.2006.11.044)
- Yagi T, Ikawa Y, Yoshida K et al (1990) Homologous recombination at c-fyn locus of mouse embryonic stem cells with use of diphtheria toxin A-fragment gene in

negative selection. *Proc Natl Acad Sci USA* 87:9918–9922. doi:[10.1073/pnas.87.24.9918](https://doi.org/10.1073/pnas.87.24.9918)

Yang Y, Seed B (2003) Site-specific gene targeting in mouse embryonic stem cells with intact bacterial artificial chromosomes. *Nat Biotechnol* 21:447–451. doi:[10.1038/nbt803](https://doi.org/10.1038/nbt803)

Yu Y, Bradley A (2001) Engineering chromosomal rearrangements in mice. *Nat Rev Genet* 2:780–790. doi:[10.1038/35093564](https://doi.org/10.1038/35093564)

Zhang H, Hasty P, Bradley A (1994) Targeting frequency for deletion vectors in embryonic stem cells. *Mol Cell Biol* 14:2404–2410