Genomics 93 (2009) 196–204

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Next generation tools for high-throughput promoter and expression analysis employing single-copy knock-ins at the *Hprt1* locus

G.S. Yang [a,1], K.G. Banks [b,1], R.J. Bonaguro [b], G. Wilson [a], L. Dreolini [a], C.N. de Leeuw [b,d], L. Liu [c], D.J. Swanson [c,2], D. Goldowitz [c,2], R.A. Holt [a,e], E.M. Simpson [b,d,*]

[a] Canada's Michael Smith Genome Sciences Centre, British Columbia Cancer Agency, Vancouver, British Columbia, Canada V5S 4Z6
[b] Centre for Molecular Medicine and Therapeutics, Child and Family Research Institute, University of British Columbia, Vancouver, British Columbia, Canada V5Z 4H4
[c] Department of Anatomy and Neurobiology, The University of Tennessee Health Science Center, Memphis, Tennessee 38163, USA
[d] Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia, Canada
[e] Department of Psychiatry, University of British Columbia, Vancouver, British Columbia, Canada

ARTICLE INFO

Article history:
Received 1 May 2008
Accepted 17 September 2008
Available online 3 December 2008

Keywords:
Embryonic stem cells
Gene targeting
Vectors
Genomics
Hprt1
Knock-in
Mice
Promoter regions
Transgenic

ABSTRACT

We have engineered a set of useful tools that facilitate targeted single copy knock-in (KI) at the hypoxanthine guanine phosphoribosyl transferase 1 (*Hprt1*) locus. We employed fine scale mapping to delineate the precise breakpoint location at the $Hprt1^{b\text{-}m3}$ locus allowing allele specific PCR assays to be established. Our suite of tools contains four targeting expression vectors and a complementing series of embryonic stem cell lines. Two of these vectors encode enhanced green fluorescent protein (EGFP) driven by the human cytomegalovirus immediate-early enhancer/modified chicken beta-actin (CAG) promoter, whereas the other two permit flexible combinations of a chosen promoter combined with a reporter and/or gene of choice. We have validated our tools as part of the Pleiades Promoter Project (http://www.pleiades.org), with the generation of brain-specific EGFP positive germline mouse strains.

Introduction

The transgenic mouse continues to play an important role in the study of human development, physiology, and disease. We have designed a collection of tools that facilitate single-copy knock-in at the selectable hypoxanthine guanine phosphoribosyl transferase 1 (*Hprt1*) locus in mouse embryonic stem cells (ESCs). Consistent levels of transgene expression and little to no phenotypic variability in mouse strains derived from knock-ins at this defined insertion site produce uniformity of results and do not require examination of large numbers of mouse strains [1]. This is in contrast to random insertion experiments, where the initial knock-in step is simpler, but often leads to undesirable effects such as uncontrolled copy number, variable expression levels, variability in phenotype, and insertional mutagenesis, necessitating the characterization of multiple strains of mice [2–5]. The *Hprt1* tool set we describe here facilitates targeted insertion such that the difficulty of the knock-in is less of a consideration in experimental design.

The HPRT1 enzyme is important in the salvage of purines for biosynthesis of nucleotides, providing an alternate to the *de novo* nucleotide synthesis pathways. Specifically, HPRT1 catalyzes the conversion of hypoxanthine and guanine to the mononucleotides IMP and GMP, respectively [6–9]. The *Hprt1* locus, located on the X Chromosome (Chr), is 44 kb in size and comprised of 9 exons [9]. HPRT1 is ubiquitously expressed in all tissues, and the *Hprt1* locus is relatively neutral as a transgene docking site because the locus has minimal influence on transgene expression when driven by an exogenous promoter [1,2,10]. Although functional deletion of *HPRT1* in humans leads to Lesch–Nyhan syndrome (reviewed in [11]), only a mild phenotype is observed in the mouse knockout [12–14], but in the knock-in (KI) mouse the sequences typically introduced into the locus complements the gene and *Hprt1* is expressed. Thus, the locus is useful in mice for single-copy gene targeting. A spontaneous deletion occurring at the mouse *Hprt1* locus, $Hprt1^{b\text{-}m3}$ [15], has been well characterized in embryonic stem cells and mice (E14TG2a, HM-1 and

* Corresponding author. Centre for Molecular Medicine and Therapeutics, Child and Family Research Institute, University of British Columbia, Vancouver, British Columbia, Canada V5Z 4H4. Fax: +1 604 875 3819.
E-mail address: simpson@cmmt.ubc.ca (E.M. Simpson).
[1] These authors contributed equally to this work.
[2] Current address: Centre for Molecular Medicine and Therapeutics, Child and Family Research Institute, and Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia, Canada V5Z 4H4.

0888-7543/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2008.09.014

B6.129P2-*Hprt1*$^{b\text{-}m3}$/J), and can be targeted using plasmids [16–18] and much larger bacterial artificial chromosome (BAC) constructs [10]. This spontaneous deletion removes two essential exons from the 5′ end of the *Hprt1* locus. Complementation of the locus by homologous recombination with KI vectors containing the two deleted exons [16] allows positive selection of recombinants. A large majority of selected clones are typically correctly targeted [19–21]. This is an important advantage over more standard positive–negative selection systems for targeting. The ability to select, both positively, with HAT (hypoxanthine, aminopterin, thymidine), and negatively, with 6-thioguanine, for targeting events at this locus establishes *Hprt1* as a valuable tool for transgenesis both *in vitro*, and *in vivo* [9,17]. The X Chr location of the *Hprt1* locus can be a concern given that single copy insertions in females are subject to random X-inactivation, which may result in mosaic expression of the transgene [10]. However, the X Chr location of the *Hprt1* locus can also be beneficial as all the daughters of a carrier father will have the transgene, thus reducing the need for genotyping.

We have advanced the existing *Hprt1* targeting and expression system [16] and developed more sophisticated tools, including expression vectors, embryonic stem cells (ESCs), and PCR assays, for mouse genome modification at this locus. A new vector, pEMS1306, has been designed to facilitate the single copy delivery of either a promoter driving EGFP and/or a gene of interest to the *Hprt1* locus. In our current application, different promoter constructs of interest are inserted into the multiple cloning site and used to drive expression of a downstream enhanced green fluorescent protein (EGFP) reporter.

We have also developed two series of new hybrid ESCs that show high germline competence. These ESCs are hemizygous for the *Hprt1*$^{b\text{-}m3}$ spontaneous deletion and heterozygous or wild-type (WT) for *Gt(ROSA)26Sor*^{tm1Sor}, which serves as a *Cre* reporter and historical marker. As with the traditional E14TG2a ESC line, these ESCs have proven useful for making new transgenic strains by conventional microinjection. An additional benefit of these new ESC lines is the ability to use the co-culture technique to produce germline competent chimeras, which is not permissible using E14TG2a [21]. These vectors and ESCs have also enabled us to derive a new strain of mice carrying the human cytomegalovirus immediate-early enhancer/modified chicken beta-actin (CAG) promoter driving the EGFP tetracysteine reporter from pEMS1306, knocked in to the *Hprt1*$^{b\text{-}m3}$ locus. Moreover we have defined the exact boundaries of the spontaneous breakpoint at *Hprt1*$^{b\text{-}m3}$ carried by E14TG2a and the mouse strain B6.129P2-*Hprt1*$^{b\text{-}m3}$, enabling us to design new specific PCR assays for this and the *Hprt1* WT alleles. We have also designed primers that are specific for the *HPRT1* complementary sequence (*HPRT1*-CS: human-to-mouse) at the *Hprt1* locus only after it is complemented by a targeting event.

Results

Multi-featured vectors enable and expand functionality of docking at the Hprt1 locus

Two targeting vectors, pEMS1306 and pEMS1307 (Fig. 1), were developed to facilitate single copy insertion of transgenes targeted to the *Hprt1* locus. The backbone sequence of both vectors contains the homologous sequence required to target the *Hprt1* locus in mouse ESCs that contain the *Hprt1*$^{b\text{-}m3}$ deleted allele. Proper integration of the vectors complements the structure and function of the *Hprt1* gene and allows positive selection of the recombinants. Our vectors encode a multiple cloning site (MCS) containing a total of five restriction sites, including the eight-cutters, FseI and AscI. These two rare restriction sites are less likely to be found within the promoter to be cloned allowing greater opportunity for directional cloning of promoter constructs with a wider range of sequence variability. The chosen promoter drives expression of a reporter gene that is an EGFP/tetracysteine motif fusion. The tetracysteine tag allows alternative

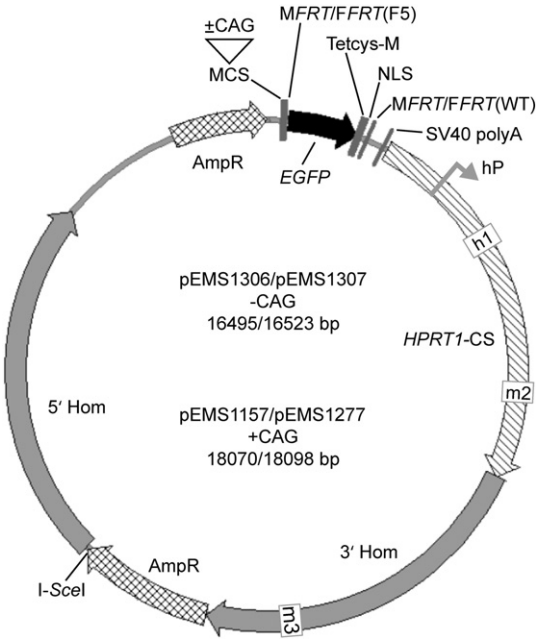

Fig. 1. Novel plasmid vectors facilitate single copy insertion at the *Hprt1* locus. pEMS1306/pEMS1307 (promoterless (−CAG)), and pEMS1157/pEMS1277 (+CAG) vectors containing *Hprt1* homology regions (5′ Hom, 3′ Hom, solid grey arrows), *HPRT1* complementary sequence (*HPRT1*-CS, slashed arrow), containing the human *HPRT1* promoter (hP), human exon 1 (h1), and mouse exon 2 (m2) and 3 (m3), multiple cloning site (MCS), minimal *FRT* (M*FRT*) and full *FRT* (F*FRT*) WT and F5 mutant sites, SV40 polyadenylation (SV40 polyA) site, nuclear localization signal (NLS), tetracysteine motif (Tetcys-M), ampicillin resistance (AmpR, checkered arrows), EGFP reporter (black arrow), and I-SceI linearization site.

detection by arsenical ligands [22]. In addition, the tetracysteine-tagged EGFP marker is flanked by 5′ mutant (F5), and 3′ wild-type (WT) *FRT* sites, allowing Flp-recombination mediated cassette exchange (RMCE) [23] of the EGFP reporter with a preferred gene of choice. The pEMS1306 and pEMS1307 targeting vectors differ only in that pEMS1306 contains minimal *FRT* (M*FRT*), 34 bp, sites compared to the full *FRT* (F*FRT*), 48 bp, version in pEMS1307 [24]. This modification was driven by the successful use of F*FRT*s for RMCE in mouse ESCs [23,25,26]. There is also an upstream Kozak consensus sequence to increase the translation efficiency of the reporter gene in eukaryotic cells in addition to three copies of a nuclear localization signal (NLS) for increased translocation efficiency of the protein into the nucleus of the cells. An SV40 polyadenylation signal directs proper processing of the mRNA 3′ end in both of these constructs.

Two additional constructs, pEMS1277, and pEM1277, were generated using the ubiquitously expressed human cytomegalovirus immediate-early enhancer/modified chicken beta-actin (CAG) promoter to test the functionality of the EGFP reporter in the pEMS1306 and pEMS1307 vector backbones. Four constructs, pEMS1306, pEMS1157, pEMS1307, and pEMS1277 were electroporated into mEMS1204 ESCs, and all of the constructs generated colonies upon HAT selection indicating proper homologous recombination and complementation of the *Hprt1*$^{b\text{-}m3}$ locus. Colonies were picked for screening by PCR to confirm the presence of the transgene immediately upstream of the *Hprt1* locus. The high frequency (90–95%) of correct homologous recombination events we observe at the *Hprt1*$^{b\text{-}m3}$ locus is similar to previously published data [10,16,18–20].

Enabling the utilization of the Hprt1$^{b\text{-}m3}$ *spontaneous deletion locus by fine scale mapping*

Although the sequences flanking the breakpoint in the E14TG2a cell line (and subclones) and *Hprt1*$^{b\text{-}m3}$ mice are characterized [27], the exact breakpoint has never been defined. We used direct sequence