Methods 121–122 (2017) 45–54

Contents lists available at ScienceDirect

Methods

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

Versatile and precise gene-targeting strategies for functional studies in mammalian cell lines

M. Wassef [a,b,c,*], A. Luscan [d,e], A. Battistella [a,b,c], S. Le Corre [a,b,c], H. Li [f], M.R. Wallace [f,g,h], M. Vidaud [d,e], R. Margueron [a,b,c]

[a] Institut Curie, PSL Research University, 75005 Paris, France
[b] INSERM U934, Paris, France
[c] CNRS UMR3215, Paris, France
[d] INSERM UMR_S745 et EA7331, Université Paris Descartes, Sorbonne Paris Cité, Facultée des Sciences Pharmaceutiques et Biologiques, 75006 Paris, France
[e] Service de Biochimie et Génétique Moléculaire, Hôpital Cochin, Assistance Publique-Hôpitaux de Paris, 75014 Paris, France
[f] Department of Molecular Genetics and Microbiology, University of Florida College of Medicine, Gainesville, FL, USA
[g] University of Florida Health Cancer Center, University of Florida, Gainesville, FL, USA
[h] University of Florida Genetics Institute, University of Florida, Gainesville, FL, USA

ARTICLE INFO

Article history:
Received 20 December 2016
Received in revised form 3 May 2017
Accepted 5 May 2017
Available online 10 May 2017

Keywords:
Genome editing
CRISPR/CAS9
Targeting vector
Reversible knockout
Constitutive knockout
Point mutation

ABSTRACT

The advent of programmable nucleases such as ZFNs, TALENs and CRISPR/Cas9 has brought the power of genetic manipulation to widely used model systems. In mammalian cells, nuclease-mediated DNA double strand break is mainly repaired through the error-prone non-homologous end-joining (NHEJ) repair pathway, eventually leading to accumulation of small deletions or insertions (indels) that can inactivate gene function. However, due to the variable size of the indels and the polyploid status of many cell lines (e.g., cancer-derived cells), obtaining a knockout usually requires lengthy screening and characterization procedures. Given the more precise type of modifications that can be introduced upon homology-directed repair (HDR), we have developed HDR-based gene-targeting strategies that greatly facilitate the process of knockout generation in cell lines. To generate reversible knockouts (R-KO), a selectable promoter-less STOP cassette is inserted in an intron, interrupting transcription. Loss-of-function can be validated by RT-qPCR and is removable, enabling subsequent restoration of gene function. A variant of the R-KO procedure can be used to introduce point mutations. To generate constitutive knockouts (C-KO), an exon is targeted, which makes use of HDR-based gene disruption together with NHEJ-induced indels on non-HDR targeted allele(s). Hence the C-KO procedure greatly facilitates simultaneous inactivation of multiple alleles. Overall these genome-editing tools offer superior precision and efficiency for functional genetic approaches. We provide detailed protocols guiding in the design of targeting vectors and in the analysis and validation of gene targeting experiments.

Contents

* Corresponding author at: Institut Curie, PSL Research University, 75005 Paris, France.
E-mail address: michel.wassef@curie.fr (M. Wassef).

http://dx.doi.org/10.1016/j.ymeth.2017.05.003
1046-2023/

1. Introduction

Genetic modification by homologous recombination (HR) between genomic DNA and an exogenously provided DNA template is considered as a gold standard for genome engineering, enabling precise modifications to be made such as knockouts, point mutations or insertion of tags [1]. Double strand break (DSB) following sequence-specific recruitment of a nuclease greatly stimulates HR by way of the homology directed repair pathway (HDR), either using a double-stranded template [2] or single stranded oligonucleotide [3]. An alternative repair pathway, based on non-homologous end-joining (NHEJ) leads to indels of variable size that can result in a frameshift of a protein-coding gene (see [4] for a review). NHEJ-repair is prevalent in most mammalian cell lines and is thus frequently used to generate knockouts, while HDR is typically used to introduce defined modifications such as insertions and point mutations [5]. These approaches have become much easier to implement with the advent of CRISPR/CAS9 technology. However, important bottlenecks remain that complicate genetic engineering in a number of model systems. First, genome editing by programmable nucleases is subject to off-target mutagenesis, which can confound the analysis of the mutant phenotype. Second, due to the variable size of NHEJ-induced indels, generating a full KO (i.e., harboring out-of-frame indels on all alleles) requires sub-cloning of the cell line of interest, which can also lead to confounding effects. In addition, many commonly used cancer cell lines are polyploid and thus harbor more than two gene copies, complicating the process of inactivating all alleles at once. Third, due to the prevalence of NHEJ over HDR in most cell lines, introduction of a defined modification by HDR is often accompanied by NHEJ-induced indels on the non-targeted allele(s). These additional events can complicate the analysis, especially when modeling the impact of a point mutation. Last, with a few exceptions, gene knockouts by NHEJ-induced indels or introduction of defined modifications using oligonucleotide-mediated HDR cannot be selected for, although several methods exist to enrich for transfected cells and to screen for mutant clones (e.g., [6]). Hence, genome editing using these strategies often involves tedious screening procedures.

In order to address these issues, we have developed genome-engineering strategies based on HDR, inspired by classical gene targeting tools. These approaches enable the generation of reversible knockouts, constitutive knockouts and point mutations. Our strategies take into account the occurrence of NHEJ-induced indels on non-HDR targeted alleles, either limiting their functional impact or taking advantage of their presence to facilitate gene disruption. In addition, by using promoter-less selection cassettes, these tools facilitate the selection of rare events, thus providing superior efficiency and precision over mainstream methods.

2. Results

2.1. Generation of reversible knockouts

Genome editing by programmable nucleases is subject to off-target mutagenesis and requires sub-cloning of the cell line of interest. Re-expression of the wild-type protein in mutant clones can help alleviate these confounding factors but the expression level of the rescuing protein needs to be tightly controlled in order to match endogenous levels. To address these issues, we developed a reversible knockout strategy. We built a promoter-less cassette (termed R-KO cassette for **r**eversible **k**nock**o**ut) inspired by existing gene-trap cassettes [7]. The absence of a promoter is expected to reduce the fraction of resistant-clones arising from random integration events [8,9]. The R-KO cassette is composed of the mouse *Engrailed 2* (*En2*) splice acceptor followed by a "ribosome skipping" T2A sequence, an antibiotic resistance gene, a stop codon and SV40 polyadenylation sequence (Fig. 1A). Hence, insertion of the cassette in an intron downstream of the initiation codon is expected to terminate target gene expression. The cassette is flanked by FRT sequences, enabling its flippase-mediated removal and subsequent re-expression of the target gene.

As a proof of principle, we targeted the 4th intron of *BAP1*, a gene frequently inactivated in multiple types of cancer, including uveal melanoma [10]. We chose the HAP1 near-haploid human cell line [11] since it only requires targeting of one allele. Out of 24 puromycin-positive clones resulting from transfection of a *BAP1* R-KO cassette and corresponding sgRNA/CAS9, 5 were positive for all 3 genotyping PCRs (Figure S1A, see further below for an explanation of the genotyping strategy). Expression of BAP1 protein was undetectable in 4 of these 5 clones (Fig. S1B). Downstream of the insertion point, *BAP1* transcript was completely lost as assessed by RNA-seq and RT-qPCR (Fig. 1B, C) and BAP1 protein was undetectable. We also targeted *ASXL1* and *ASXL2* genes encoding partners of BAP1. As shown in Fig. S2, insertion of the R-KO cassette within these genes similarly resulted in an acute transcription arrest downstream of the insertion point.

We then removed the cassette by transient transfection of a flippase-expressing vector. Removal of the R-KO cassette successfully restored *BAP1* RNA (Fig. 1C) and protein (Fig. 1D) expression.

Several features make the reversible knockout strategy particularly useful. First, gene function can be restored to rule out off-target and/or subclonal effects in the observed phenotype(s). Second, knockout and rescue can be assessed by reverse transcription followed by quantitative PCR (RT-qPCR), which constitutes a specific and sensitive validation procedure that exempts from having specific antibodies against the protein of interest (e.g. to validate the knockout by western blot). Third, interruption of transcription by the R-KO cassette precludes translation re-initiation down-