Apr 01, 2020

# Basic Protocol 2: Tagging a gene of interest with AID

In 1 collection

Kizhakke Mattada Sathyan[1], Thomas G. Scott[1], Michael J. Guertin[1,2,3]

[1]Biochemistry and Molecular Genetics Department, University of Virginia, [2]Center for Public Health Genomics, University of Virginia, [3]Cancer Center, University of Virginia

1 Works for me dx.doi.org/10.17504/protocols.io.bdyei7te

Michael Guertin

ABSTRACT

The next step in developing the ARF-AID system is to tag the gene of interest with AID. The ARF-AID system requires full-length AID (Figure 1) because the characterized interaction domains of AID with ARF are domains III and IV. Domains I and II are involved in the interaction with TIR1. The mini-AID lacks domains III and IV and will not interact with ARF to stabilize the protein in the absence of auxin (Sathyan et al., 2019). Note that the antibiotic selection marker (HygroR) is co-transcribed with AID and the protein products are separated during translation. Therefore, the resistance marker will be expressed at levels comparable to the target protein.

ATTACHMENTS

ARF-AID_methods.pdf

GUIDELINES

Please refer to the Guidelines & Warnings section of the protocol collection for additional guidelines.

**Design gene-specific sgRNA and cloning**

Adding any tag to the N-terminal or C-terminal of the protein could disrupt the function of a gene. Therefore, the functionality of the AID-tagged proteins should be empirically determined. We recommend designing sgRNAs to both ends and then testing the functionality of both the N and C-terminally tagged proteins. There are several free tools available to design sgRNA, such as from Benchling, CHOPCHOP, E-CRISP, and CRISPOR (Labun et al., 2019; Heigwer et al., 2014; Benchling, 2019; Concordet and Haeussler, 2018). We outline the process of using Benchling or manually choosing sgRNAs based on the presence of a PAM sequence within 25 bases of the start or stop codon.

Consider two parameters while choosing an sgRNA: the distance from the desired homologous repair site and the specificity and off-target effects of the sgRNA. The efficiency of the homologous repair of the AID-tag at the cut site increases if the required homologous repair site is near the sgRNA cut site (O'Brien et al., 2019; Inui et al., 2014). Increased distance between the cut site and the start codon or the stop codon can create challenges when designing homology arms. For example, when inserting an N-terminal tag and using a guide with a cut site upstream of the start codon, the downstream homology arm can only extend to the start codon. The upstream homology arm can either end at the cut site, removing the 5' UTR from the resulting product, or continue to the start codon. If this homology arm has extensive homology with both sides of the cut site, then the cut may be repaired without proper insertion of the template. The same types of challenges occur with cut sites internal to the protein coding region, as well as with C-terminal tagging. However, silent mutations can be introduced into the homology arms to decrease sequence homology within the protein coding region without reestablishing a sgRNA recognition site.

The sgRNA's proximity to the homologous recombination site takes priority over the specificity scores. Design at least three sgRNA to each terminus and clone into pX458 (Ran et al., 2013) (or an appropriate vector), which harbors GFP that can be used to quantify the efficiency of transfection. If an ideal sgRNA is not found, check the possibility of using other Cas9 enzymes with different PAM sequence requirements (Kleinstiver et al., 2015). Guide RNAs must be designed with the 3' PAM sequence but do not include the PAM sequence in the cloned sgRNA construct.

The U6 promoter in the pSpCas9(BB)-2A-GFP (pX458) plasmid requires a 'G' at the beginning of the guide to efficiently transcribe the sgRNA. If the sgRNA designed does not have a 'G' at the 5'-end, add one 'G' at the 5' of the forward sequence of the guide RNA and



**Citation:** Kizhakke Mattada Sathyan, Thomas G. Scott, Michael J. Guertin (04/01/2020). Basic Protocol 2: Tagging a gene of interest with AID. https://dx.doi.org/10.17504/protocols.io.bdyei7te

reverse complement of the ‘G’ in the reverse sequence. A general strategy to clone and test the efficiency of gene-specific sgRNA is shown in Figure 6.

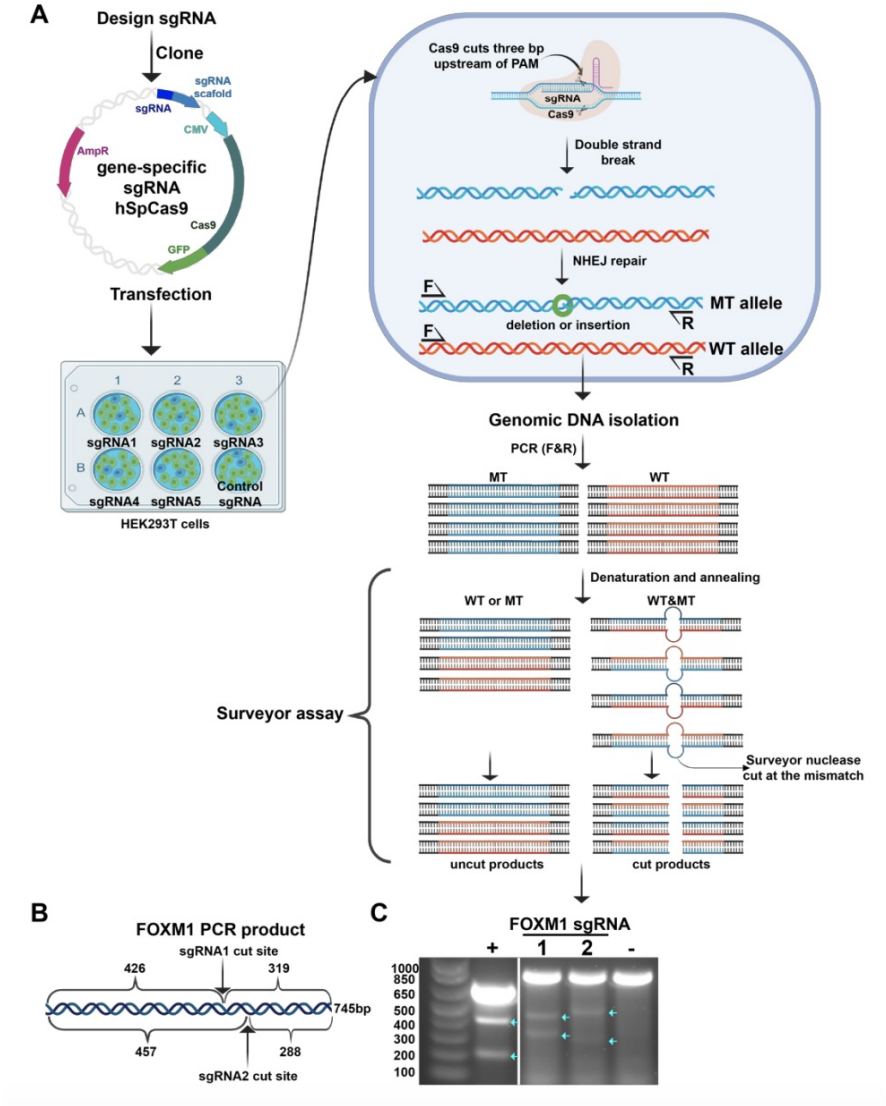


Figure 6. Design and test gene-specific sgRNA. A) A Surveyor assay is used to detect the mutation generated by non-homologous end joining (green circle) at the sgRNA targeted sites. B) Two sgRNAs were designed proximal to the 3' end of FOXM1 gene. PCR primers flank the sgRNA cutting sites and the sizes of Surveyor-cleaved products are indicated. C) The PCR product from two FOXM1 sgRNA transfected HEK293T cells were digested with surveyor and visualized. The positive control is provided in the surveyor assay kit, and the negative control is a PCR product from HEK293T cells without an sgRNA transfection. Cyan arrows indicate the bands that result from heteroduplex digestion.

### Homology directed repair construct design

Homology-Directed Repair (HDR) is the mechanism by which AID is translationally fused to the N or C terminus of the target gene. CRISPR-Cas9 is directed to the region by sgRNA and this complex cleaves double-stranded DNA, which can be repaired by non-homologous end joining or homologous recombination/repair. The presence of a repair construct is necessary to increase the probability of homologous repair. We designed repair constructs specific for the N-terminal and C-terminal of the protein (Figure 7). For C-terminal fusion, the repair construct consists of AID separated by a porcine teschovirus-1 ribosomal skipping sequence (P2A) (Kim et al., 2011) from the hygromycin resistance gene. In the N-terminal repair construct, the order is reversed: HygR-P2A-AID. In both cases, the AID is separated from the protein of interest by adding a linker sequence of 6-9 amino acids (3x GGS).



**Citation:** Kizhakke Mattada Sathyan, Thomas G. Scott, Michael J. Guertin (04/01/2020). Basic Protocol 2: Tagging a gene of interest with AID. https://dx.doi.org/10.17504/protocols.io.bdyei7te

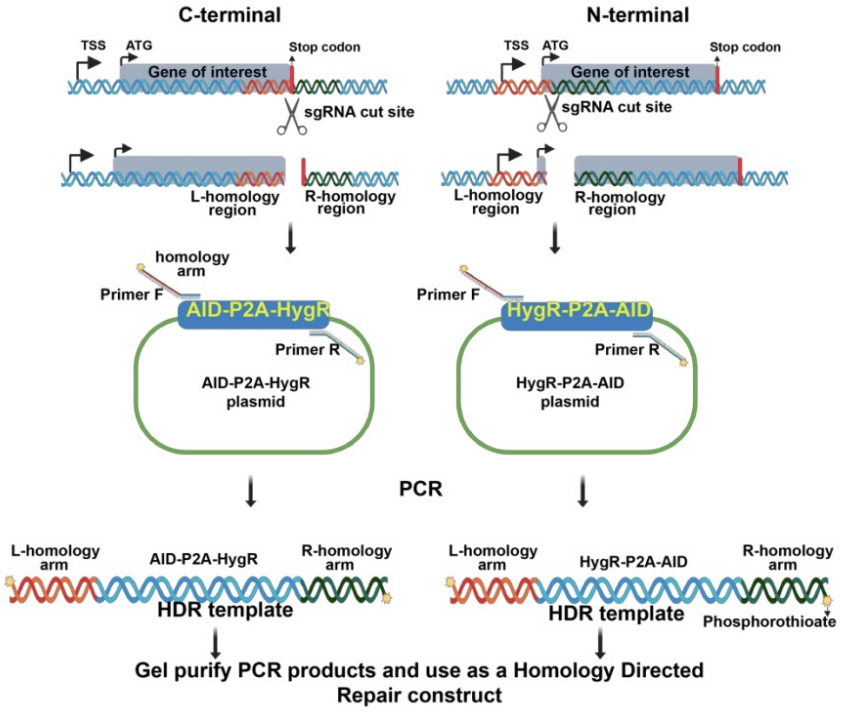


Figure 7. The strategy for designing the Homology Directed Repair (HDR) construct for both the N and C termini. A 50 nucleotide homology arm tail is added to the primers that amplify the HygR-P2A-AID (N) and AID-P2A-HygR (C) cassettes. PCR products are gel purified and used as an HDR template

The HDR construct can be PCR products or cloned into a plasmid. If PCR products are used, two phosphorothioate moieties are added to the first two 5' nucleotides of each primer to increase PCR product stability in the cell (Zheng et al., 2014). Homology arms can vary in length from eight hundred nucleotides to less than 10 nucleotides (Sakuma et al., 2016; Lambrus et al., 2018; Paix et al., 2017). We recommend 50 nucleotide homology arms on both sides of the cut site for the AID integration (Sathyan et al., 2019).

In the C-terminal region, if the cut site is before the stop codon, we include wobble substitutions for C-terminal amino acids in the forward primer. These wobble substitutions have a two-fold purpose: 1) they ensure that the only homologous region in the donor is upstream of the cut site and 2) they prevent reconstitution of the sgRNA recognition site. Similarly, in the N-terminal region, if the cut site is after the start codon, we include wobble substitutions for N-terminal amino acids in the reverse primer. Simulate HDR reaction by in silico PCR using the SnapGene software or other software. Confirm that the sgRNA recognition sequence is not recreated after the HDR and only the desired homologous sequences are available for recombination.

### Designing Primers for the C-terminal tagging of the protein

The 3' UTR can be critical in the regulation of the gene expression, and any changes in the sequence could modulate expression levels. If possible, select sgRNAs that cut inside the gene before the stop codon. If the sgRNA cuts after the stop codon, choose the closest to the stop codon. This reduces the challenges affecting the regulatory elements of the gene while designing the repair construct. Consider using another CRISPR enzyme with a different PAM recognition sequence if there is no optimal sgRNA sequence available for the CRISPR-Cas9 system (Kleinstiver et al., 2015).

### Designing homology arm primers

In all cases in which the desired homologous recombination event recreates the original guide sequence with fewer than two mismatches and an intact PAM sequence, mutate the relevant homology arm to abrogate guide binding. Use a silent mutation to destroy the PAM sequence if possible, or use two silent mutations near the 3' end of the guide sequence (Cong et al., 2013). Check the evolutionary conservation of the wobble nucleotides (Ramani et al., 2019) to prioritize less conserved nucleotides. Check a codon usage chart to prioritize codons used at a similar frequency to the replaced codon (Athey et al., 2017).

- **Upstream homology arm (coding strand primer design)**



**Citation:** Kizhakke Mattada Sathyan, Thomas G. Scott, Michael J. Guertin (04/01/2020). Basic Protocol 2: Tagging a gene of interest with AID. https://dx.doi.org/10.17504/protocols.io.bdyei7te

***Cut site upstream of the stop codon***

The upstream homology arm begins 50 bases upstream of the cut site and the last base is the nucleotide immediately upstream of the first stop codon base. The coding nucleotides downstream of the cut site will need to be modified at their wobble bases.

Append the sequence given below so that the primer anneals to the AIDP2A-HygR cassette in pMGS54.

*Critical: confirm that the AID-P2A-HygR cassette is in frame with the protein after repair by in silico PCR using SnapGene or any other program.*

***Cut site downstream of the stop codon***

The upstream homology arm begins 50 bases upstream of the cut site and the last base is the nucleotide immediately upstream of the first stop codon base.

Append 5'- GGTGGATCTGGAGGTTCAGGTGGCAGTGTCGAGCTGAATCT-3' to the 3'-end of the upstream homology arm for C-terminal tagging using the insert from pMGS54. This sequence contains a flexible linker region prior to the AID coding sequence.

*Critical: confirm that the AID-P2A-HygR cassette is in frame with the protein after repair by in silico PCR using the SnapGene or any other program.*

- **Downstream homology arm (template strand primer design)**

The downstream homology arm begins at the cut site and extends 50 bases downstream of the stop codon.

In the case of sgRNA cut site downstream of stop codon, the homology arm can extend to immediately upstream of the stop codon. This will include the full 3' UTR, but may decrease the efficiency of HDR.

Append 5'-TCAGTTAGCCTCCCCCATCTC-3' to the 3'-end of the downstream homology arm for C-terminal tagging using the insert from pMGS54. This sequence contains the template strand of the HygR coding sequence.

Phosphorothioate moieties are added to the 5' end of both upstream and downstream primers. PCR with the above primers using pMGS54 as template produces an amplicon of 1791 plus the length of the homology arm.

**Designing Primers for N-terminal tagging of the protein**

The 5' UTR can be important in the regulation of gene expression, therefore any changes in the sequence should be avoided. Select sgRNAs that cut inside the gene after or very proximal to the start codon.

- **Upstream homology arm (coding strand primer design)**

***Cut site downstream of the start codon***

Start the homology arm 50 bases upstream of the cut site and end at the cut site. If the cut site is in the middle of a codon, ensure the HygR gene is in frame with the protein by adding extra nucleotides.

Append the sequence given below so that the primer anneals to the HygR-P2A-AID cassette in pMGS58.

***Cut site upstream of the start codon***

Start the upstream homology arm 50 bases upstream of the cut site and end at the nucleotide immediately preceding the start codon. Make necessary changes in the PAM sequence or the targeting sequence to avoid repeated cutting by the repaired insertion. However, avoid these changes if possible to mitigate the risk of altering protein expression.

Append the sequence given below so that the primer anneals to the HygR-P2A-AID cassette in pMGS58.



**Citation:** Kizhakke Mattada Sathyan, Thomas G. Scott, Michael J. Guertin (04/01/2020). Basic Protocol 2: Tagging a gene of interest with AID. https://dx.doi.org/10.17504/protocols.io.bdyei7te

Append 5'-ATGAAAAAGCCTGAACTCACCG-3' to the 3'-end of the upstream homology arm for N-terminal tagging using the insert from pMGS58. This sequence contains the beginning of the HygR coding sequence.

- **Downstream homology arm (template strand primer design)**

***Cut site downstream of the start codon***

The downstream homology arm begins 50 bases downstream of the cut site and the last base is the nucleotide immediately downstream of the last start codon base. If necessary, change the wobble nucleotides of the codons before the cut site.

Append the sequence given below so that the primer anneals to the HygR-P2A-AID casette in pMGS58.

*Critical: confirm that the HygR-P2A-AID cassette is in frame with the protein after repair by in silico PCR using the SnapGene or any other program.*

***Cut site upstream of the start codon***

The downstream homology arm begins 50 bases downstream of the cut site and the last base is the nucleotide immediately downstream of the last start codon base.

To functionally separate the protein of interest from AID, a linker of 9 amino acids is added at the C-terminus of the AID in the pMGS58 plasmid. The provided primer (below) amplifies both the linker and the AID. If any other template is used for generating the tag, be sure to add linker amino acid sequence.

*Critical: confirm that the HygR-P2A-AID cassette is in frame with the protein after repair by in silico PCR using the SnapGene or any other program.*

Append 5'-CCCACCTGAACCTCCAGATC-3' to the 3'-end of the downstream homology arm for N-terminal tagging using the insert from pMGS58. This sequence is complementary to the coding sequence of a flexible linker sequence following the end of the AID coding sequence in the plasmid. The PCR with the above primers using pMGS58 as template produces an amplicon of 1815 base pairs plus the length of the homology arm.

Add phosphorothioate moieties to the first two 5' nucleotides of both upstream and downstream primers.

**AID-tagging the gene of interest**

The three main steps in the tagging of a gene with AID include: 1) cotransfection of a gene-specific sgRNA and HDR template into ARF-TIR1 progenitor cells; 2) selecting tagged clones with hygromycin B; and 3) clonal expansion and confirmation of tagging. A general outline of these steps are illustrated in Figure 8.



**Citation:** Kizhakke Mattada Sathyan, Thomas G. Scott, Michael J. Guertin (04/01/2020). Basic Protocol 2: Tagging a gene of interest with AID. https://dx.doi.org/10.17504/protocols.io.bdyei7te

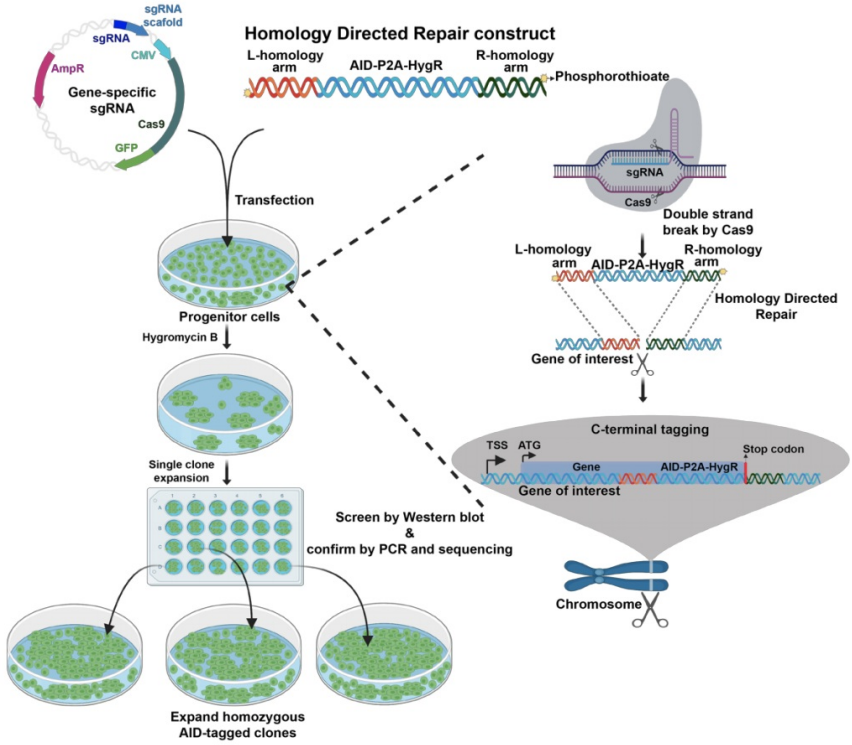


Figure 8. HDR-mediated integration of AID-tag at the 3' end of the gene of interest using CRISPRCas9. HEK293T-eGFP-ARF-P2A-TIR1 cells are cotransfected with an sgRNA and a PCR amplified repair construct. The cells are selected with hygromycin B and the clones are screened for integration using Western blotting, PCR, and sequencing.

MATERIALS

| NAME | CATALOG # | VENDOR |
|---|---|---|
| Carbenicillin (Disodium) | C-103 | Gold Biotechnology |
| Betaine solution (5M PCR Reagent) | B0300 | Sigma – Aldrich |
| QIAquick Gel Extraction Kit | 28704 | Qiagen |
| T4 DNA ligase | | New England Biolabs |
| Platinum™ Taq DNA Polymerase | 10966034 | Thermo Fisher |
| S.O.C. Medium | 15544034 | Thermo Fisher |
| BSA | B9000S | New England Biolabs |
| HEK293T ARF-P2A-TIR1 progenitor cells | | |
| pMGS54 (AID-P2A-Hygromycin) | 126583 | addgene |
| pMGS58 (Hygromycin-P2A-AID) | 135311 | addgene |
| BbsI | R0539S | New England Biolabs |
| T4 Polynucleotide kinase | M0201S | New England Biolabs |
| Max efficiency DH5a | 18258-012 | Invitrogen - Thermo Fisher |
| Surveyor kit | 706025 | IDT |



**Citation:** Kizhakke Mattada Sathyan, Thomas G. Scott, Michael J. Guertin (04/01/2020). Basic Protocol 2: Tagging a gene of interest with AID. https://dx.doi.org/10.17504/protocols.io.bdyei7te

| NAME | CATALOG # | VENDOR |
|---|---|---|
| LB plate | | |
| Hygromycin B | 10687010 | Invitrogen - Thermo Fisher |
| LB broth Media | BP1426 | Fisher Scientific |
| LB agar | BP1425 | Fisher Scientific |
| QIAprep Spin Miniprep Kit Print | 27104 | Qiagen |

MATERIALS TEXT

**Additional Materials:**

- sgRNA sequencing primers
  LKO-1 5' primer 5'-GACTATCATATGCTTACCGT-3'
  U6 promoter primer 5'-CACAAAGATATTAGTACAAAATACG-3'

FOXM1 sgRNA1 5'-GCAGGGCTCTACTGTAGCTC-3'
FOXM1 sgRNA2 5'-GGGACCAGTTGATGTTGTCA-3'
FOXM1 forward primer 5'-TCTGGCAGTCTCTGGATAATGAT-3'
FOXM1 reverse primer 5'-GCTGATGGATCTCAGCACCACTC-3'

SAFETY WARNINGS

Please refer to the Safety Data Sheets (SDS) for safety and environmental hazards.

BEFORE STARTING

Please refer to the Guidelines & Warnings section of the protocol collection for additional information.

### Designing sgRNA using Benchling

- **Importing target gene sequence**

1. From the left navigation bar in Benchling, click Create > CRISPR > CRISPR Guides.
2. Search for the target gene by gene ID or name, and choose the appropriate genome assembly and transcript. Optional: if there are too few nucleotides imported upstream of the start codon or downstream of the stop codon for the desired homology arm length, choose “Show Advanced Options” to import additional nucleotides.
3. Guide parameters can be left at the default settings or adjusted as needed.

- **Choosing guides**

1. Highlight the region 25 nucleotides upstream and 25 nucleotides downstream around the start codon (for N-terminal tagging) or the stop codon (for C-terminal tagging) and click “Create” to create a target sequence.
2. Benchling provides a list of guides targeting the region, along with their predicted On-Target Scores (Doench et al., 2016) and Off-Target Scores (Hsu et al., 2013). In addition to maximizing these scores, distance to the start or stop codon should be minimized (see above).
3. For guides chosen, click “Assemble” and choose the vector into which they will be cloned, e.g. pX458.
4. For each guide, the resulting Assembly shows the predicted plasmid after cloning.
5. Return to the tab with the imported sequence, click “copy the primer list”, and paste the sequences for all generated oligonucleotides into a spreadsheet.

*The two oligonucleotides (FWD and REV) can be synthesized for downstream cloning.*

Alternatively, choose the 20 nucleotide target sequence without the PAM and add overhangs to clone into pX458. If the sequence does not start with a G, add a G to the 5'-end. Append 5'-CACC-3' to the 5'-end of the forward sequence and 5'-AAAC-3' to the 5'-end of



**Citation:** Kizhakke Mattada Sathyan, Thomas G. Scott, Michael J. Guertin (04/01/2020). Basic Protocol 2: Tagging a gene of interest with AID. https://dx.doi.org/10.17504/protocols.io.bdyei7te

the reverse complement of the target sequence, ensuring that the 3' base is the complementing C nucleotide (Cong et al., 2013).

## Cloning the guide RNA into pSpCas9(BB)-2A-GFP (pX458) – Digestion of the vector with BbsI

1 Digest **3 µg** of pX458 with BbsI **Overnight** at **37 °C**

*The overnight digestion with BbsI significantly reduces negative colonies.*

**BbsI reaction mix:**

| | |
|---|---|
| NEB buffer 2.1 | 5.0 µl |
| BbsI enzyme | 1.0 µl |
| 100X BSA | 0.5 µl |
| pX458 plasmid | 3.0 µg |
| Water | to 50 µl |

## Gel purification of the digested plasmid

2 Add **8 µl** loading dye to the mixture and load onto **1 %** agarose gel. Use the same amount of the undigested pX458 vector as a negative control.

3 Use a transilluminator to excise the digested band using a clean scalpel and use the QIAquick gel extraction kit (Qiagen, 28704) to isolate the digested plasmid.

*The use of the blue light dark reader transilluminator (Clare Chemical Research, USA) may reduce DNA damage during gel excision. The digested plasmid runs between supercoiled and relaxed DNA in the undigested control sample.*

## DNA purification

4

*The method described here is adapted from the kit manual with minor changes (QIAquick gel extraction kit, Qiagen, 28704).*

Add three volumes of buffer QG to one volume of gel; consider **1 mg** of gel equivalent to **1 µl** . For example, add **300 µl** of QG buffer for **100 mg** of gel.

5

Incubate at **50 °C** for **00:10:00** . Mix by vortexing every **00:03:00** to dissolve the gel. Once the gel has completely dissolved, add **10 µl** of **3 Molarity (M)** sodium acetate, **pH5** , irrespective of the color of the mixture, and briefly mix by vortexing.

*The addition of sodium acetate significantly increases the final yield of the DNA.*



**Citation:** Kizhakke Mattada Sathyan, Thomas G. Scott, Michael J. Guertin (04/01/2020). Basic Protocol 2: Tagging a gene of interest with AID. https://dx.doi.org/10.17504/protocols.io.bdyei7te

6

Add one gel-volume of isopropanol and vortex briefly.

7

Transfer the mixture into the QIAquick column and centrifuge at **17000 x g** for **00:01:00** .

*If there is more dissolved gel mixture left, add to the same column and repeat step 7.*

8

Remove the flow-through and place the column back to the collection tube. Add **500 µl** QG buffer to the column and centrifuge at **17000 x g** for **00:01:00** .

9

Remove the flow-through and place the column back to the collection tube. To wash the column, add **750 µl** of PE buffer and centrifuge at **17000 x g** for **00:01:00** .

10

Discard flow-through and replace the column into the same tube and centrifuge again at **17000 x g** for **00:02:00** .

11

Place the column into a new clean 1.5 ml Eppendorf tube and add **50 µl** DNase and RNase free water. Incubate for **00:02:00** and centrifuge at **9000 x g** for **00:02:00** .

*Centrifuging at 9000 x g reduces the chance of breaking off the Eppendorf tube's lid.*

12 Quantify DNA using a NanoDrop and store at **-20 °C**

## Ligation of guide sequence into pX458 plasmid

13 Make **100 Micromolar (µM)** solution of the forward and reverse strands of the sgRNA guide sequence in nuclease free water.



**Citation:** Kizhakke Mattada Sathyan, Thomas G. Scott, Michael J. Guertin (04/01/2020). Basic Protocol 2: Tagging a gene of interest with AID. https://dx.doi.org/10.17504/protocols.io.bdyei7te

14 Phosphorylate and anneal the forward and reverse oligonucleotides of the sgRNA guide sequence in one reaction.

**Phosphorylation and annealing reaction mix:**

| | |
|---|---|
| Forward (100 µM) | 1 µl |
| Reverse (100 µM) | 1 µl |
| T4 DNA Ligase buffer | 1 µl |
| T4 PNK | 1 µl |
| Water | 6 µl |

Incubate in a thermocycler

**37 °C** **00:30:00**

**95 °C** **00:05:00**

Ramp down to **4 °C** at **4 °C** /min rate.

15 Dilute the annealed sgRNA guide sequence pair 20x by adding **190 µl** of nuclease-free water. We use one microliter of this **500 Nanomolar (nM)** dsDNA product for ligation.

**DNA ligation mix:**

| | |
|---|---|
| BbsI digested pX458 | 50 ng |
| annealed sgRNA | 1 µl |
| T4 ligase Buffer | 1 µl |
| T4 DNA Ligase | 1 µl |
| Water | to 10 µl |

16

Incubate ligation mix for **02:00:00** at **Room temperature** or **16 °C** **Overnight** . Include no insert control to see the rate of negative colonies.

*Typically, we incubate 2 hours at room temperature and start the transformation. Keep the rest of the ligation mix at 16oC overnight. If the first transformation does not result in colonies, then retransform the ligated product the next day.*

17 Add **2 µl** of the ligated product to **20 µl** chemically competent E. coli. We used Max efficiency DH5a competent cells from Invitrogen for transformation.

18

Incubate **On ice** for **00:30:00** and then heat shock for **00:00:30** at **42 °C** .

19 Put back **On ice** for **00:02:00** .

20

Add **250 µl** LB or SOC media and incubate at **37 °C** for **01:00:00**



**Citation:** Kizhakke Mattada Sathyan, Thomas G. Scott, Michael J. Guertin (04/01/2020). Basic Protocol 2: Tagging a gene of interest with AID. https://dx.doi.org/10.17504/protocols.io.bdyei7te

21 Plate the cells ( **272 µl** ) on a carbenicillin plate (stable version of ampicillin) and incubate at **37 °C** **Overnight** .

22 Pick three colonies from each plate and inoculate in LB containing carbenicillin and incubate **Overnight** at **37 °C** .

23 Pellet the bacterial culture by centrifugation at **17000 x g** for **00:01:00** and follow the steps in the QIAprep spin miniprep kit. Elute DNA in **50 µl** nuclease free water.

24 Confirm the sgRNA insertion by sequencing using the LKO.1 5' primer or the U6 promoter primer.

## Testing sgRNA using surveyor assay – Transfection of the sgRNA construct

25 Plate 30% confluent HEK293T cells in 6 well plate (start with 2 x $10^5$ cells).

26 Transfect **1 µg** of sgRNA plasmid with lipofectamine 3000 reagents as described in steps 5 – 9 from basic protocol 1.

27 **Twenty-four hours after transfection**, replace media with **2 ml** new media.

*GFP in the pX458 plasmid allows easy assessment of the transfection efficiency. Observe the transfected cells under a fluorescent microscope using a green filter. Higher transfection rates make it easier to assess the efficacy of the sgRNA using surveyor assay*

28 **Seventy-two hours after transfection**, remove **1 ml** of media and collect the cells in the remaining **1 ml** of media by pipetting up and down and transfer into a 1.5 ml tube.

29 Centrifuge cells at **6000 x g** for **00:02:00** using a fixed angle rotor table top centrifuge and remove media. Cells can be stored at **-20 °C** or directly begin genomic DNA isolation.

30 Proceed with genomic DNA isolation, as described in steps 25 – 35 from basic protocol 1.



**Citation:** Kizhakke Mattada Sathyan, Thomas G. Scott, Michael J. Guertin (04/01/2020). Basic Protocol 2: Tagging a gene of interest with AID. https://dx.doi.org/10.17504/protocols.io.bdyei7te

Genomic DNA PCR – Primer design

31

Make a PCR master mix by adding all the components except genomic DNA and gene specific primers adjusted for the total number of reactions.

*To amplify the region targeted by the sgRNA, design a set of forward and reverse primers. The forward primer should be ~250 bp upstream from the sgRNA cutting site and the reverse primer should be ~250 bp downstream. To facilitate mutation detection, we recommend an amplicon size of 500bp to 1Kb with an sgRNA cut site close to the center of the amplicon. There are several tools to pick appropriate primer sets, such as from Primer3, IDT or Benchling (Untergasser et al., 2012; Benchling, 2019) https://www.idtdna.com/pages/tools/primerquest). The double stranded DNA breaks generated by CRISPR-Cas9 are typically repaired by error prone non-homologous repair, which results in several types of mutations. The mutated PCR products, when annealed with the wild type or other mutated PCR products, generate a mismatch proximal to the cut site. The surveyor enzymes recognize this mismatch and cleave the heteroduplex at the site of mismatch (Figure 6). The assay produces three fragments, one undigested and two digested fragments. Depending on the site of mismatch, the digested fragments run either as one or two bands (Figure 6B&C). If the sgRNA cutting site is at the center of the PCR product, it generates two unresolvable fragments, whereas unequal fragments run as distinct bands. There is no need to mix wild type PCR products with the PCR products of the sgRNA transfected cells for the assay because many different repair products will form and other sites in the population will remain unmodified. One set of primers is required for each terminus of the gene to test sgRNAs that target these sites.*

*We used Platinum Taq DNA polymerase or Platinum Taq DNA Polymerase High Fidelity for genomic PCR. For each primer set, it is necessary to determine the annealing temperature empirically. We find that an annealing temperature within 3 degrees of the lowest melting temperature (Tm) works well. Additionally, adding denaturants such as DMSO or Betaine may help to amplify GC rich genomic regions.*

32

Aliquot **18 µl** of the master mix into 0.2 ml PCR tubes and add **5 µl** of [M]**10 ng/µl** genomic DNA into each reaction mix. The parental HEK293T DNA serves as a negative control.

33 Add **1 µl** each of forward and reverse primers of the corresponding gene targeted.



**Citation:** Kizhakke Mattada Sathyan, Thomas G. Scott, Michael J. Guertin (04/01/2020). Basic Protocol 2: Tagging a gene of interest with AID. https://dx.doi.org/10.17504/protocols.io.bdyei7te

34

**PCR mix:**

| | |
|---|---|
| Genomic DNA | 5.0 µl |
| Gene Primer F | 1.0 µl |
| Gene Primer R | 1.0 µl |
| 10X HF buffer | 2.5 µl |
| DMSO (100%) | 0.5 µl |
| 10 mM DNTP | 1.0 µl |
| MgSO4 | 1.0 µl |
| Platinum Taq DNA Polymerase High Fidelity | 0.5 µl |
| Water | 12.5 µl |

*Annealing and extension of the PCR steps depends on the primers and Taq DNA polymerase used.*

**Run PCR as follows:**
Initial denaturation

**95 °C** **00:05:00**

30 cycles of

**95 °C** **00:00:30**

___°C. **00:00:30** (Annealing temperature changes based on the primers used)

**68 °C** **00:01:00**

Final extension

**68 °C** **00:10:00**

**4 °C** hold ∞

35

Check PCR amplification by running **5 µl** of the products on an agarose gel. Add **2 µl** 6x DNA sample buffer to **5 µl** of PCR product and load onto a [M]**1 %** agarose gel.

*Keep the remainder of the PCR product for the surveyor assay. This step makes sure the PCR worked before starting the surveyor assay.*

36

Run the samples at constant 90V and visualize the bands using a UV transilluminator (follow steps 38 – 42 from basic protocol 1). If there is only one bright PCR product, proceed with the surveyor assay.



**Citation:** Kizhakke Mattada Sathyan, Thomas G. Scott, Michael J. Guertin (04/01/2020). Basic Protocol 2: Tagging a gene of interest with AID. https://dx.doi.org/10.17504/protocols.io.bdyei7te

Surveyor assay

37

*This step involves denaturation followed by annealing of the PCR products to form heteroduplexes, followed by a surveyor nuclease reaction. The denaturation and renaturation step is essential because the final cycle of PCR generates homoduplexes that are not recognized by the surveyor nuclease.*

Denature the PCR products from **go to step #34** at **95 °C** for **00:10:00** and then allow to renature stepwise using the following program:

**95 °C** **00:10:00**
**85 °C** **00:01:00**
**75 °C** **00:01:00**
**65 °C** **00:01:00**
**55 °C** **00:01:00**
**45 °C** **00:01:00**
**35 °C** **00:01:00**
**25 °C** **00:01:00**
**4 °C** hold ∞

A ramp down rate of **0.3 °C** /sec is recommended.

38 **Surveyor nuclease reaction mix:**

| | |
|---|---|
| Reannealed PCR product | 20.0 µl |
| Surveyor Nuclease S | 1.0 µl |
| Surveyor Enhancer S | 1.0 µl |
| 0.15 M MgCl2 Solution | 2.0 µl |

39

Mix by pipetting and incubate at **42 °C** for **01:00:00** in a thermocycler.

40

Stop the reaction by adding **2.4 µl** of Stop Solution and mix.

*Either directly electrophoresis the sample on a 2% agarose gel after mixing with loading dye or store at -20°C for future use.*



**Citation:** Kizhakke Mattada Sathyan, Thomas G. Scott, Michael J. Guertin (04/01/2020). Basic Protocol 2: Tagging a gene of interest with AID. https://dx.doi.org/10.17504/protocols.io.bdyei7te

41

Run the samples at constant 90V and visualize the bands using a UV transilluminator. Follow steps 38 – 42 from basic protocol 1 for running and visualizing the samples. The ratio of the undigested band with that of the digested band gives an estimate of relative sgRNA efficiency.

*As an example, we designed two sgRNA that target the 3' end of FOXM1 coding regions (Figure 6B&C).*

PCR amplification of Homology Directed Repair construct

42 Synthesize the designed primers from Integrated DNA Technologies (IDT), adding phosphorothioate as a modification to the primer in the details.

43

Amplify the HDR template using the primers and Platinum Taq DNA Polymerase High Fidelity. We perform several 50 µl reactions (typically 4 to 8), and gel purify the PCR products.

**PCR mix:**

| | |
|---|---|
| Plasmid DNA | 50.0 ng |
| 10 µM Primer F | 1.0 µl |
| 10 µM Primer R | 1.0 µl |
| 10XHigh Fidelity buffer | 5.0 µl |
| DMSO | 0.5 µl |
| 10 mM DNTP | 2.0 µl |
| MgSO4 | 2.0 µl |
| Platinum Taq DNA Polymerase High Fidelity | 0.5 µl |
| Water | to 50 µl |

**Use PCR condition as follows:**

Initial denaturation

**95 °C** **00:05:00**

30 cycles of

**95 °C** **00:00:30**

**60 °C** **00:00:30**

**68 °C** **00:01:00**

Final extension

**68 °C** **00:10:00**

**4 °C** hold ∞

*Removing any remaining primer from the PCR products is important as primers may interfere with homologous recombination. The primers are very long, so conventional PCR clean-up kits will not remove primers efficiently and reduce the HDR efficiency by binding to the cut site. Always gel purify the PCR products.*



**Citation:** Kizhakke Mattada Sathyan, Thomas G. Scott, Michael J. Guertin (04/01/2020). Basic Protocol 2: Tagging a gene of interest with AID. https://dx.doi.org/10.17504/protocols.io.bdyei7te

## Agarose gel purification

44 Run the PCR products on an agarose gel and cut out the repair construct band.

45 Combine all the gel slices into one tube and purify the DNA using the Qiagen gel purification kit similar to steps 2 – 12.

46 Elute DNA using **50 µl** to **100 µl** nuclease-free water and quantify using a NanoDrop and store at **-20 °C** . DNA is stable in **-20 °C** for several months.

## Transfection

47 Grow the ARF-AID progenitor cells and split into six-well plates to ~30% confluency.

48 Cotransfect **1 µg** of gene-specific sgRNA plasmid and **400 ng** of double-stranded homology repair template PCR product to the cells as described in steps 1 – 9 of basic protocol 1. Use **5 µl** – **7.5 µl** Lipofectamine reagent for transfection. The parental pX458 cotransfected with HDR PCR product is used as a transfection control and we recommend keeping a untransfected control cell well.

*Transfect multiple (at least 4 wells) wells to get enough colonies.*

49 Replace with new media **24 hours after transfection**.

50 **After forty-eight hours**, expand each well into a 10 cm plate.

*Expanding cells into 10 cm plates spreads cells so they form isolated colonies.*

51 **Seventy-two hours after transfection**, add [M]**100 µg/ml** final concentration of hygromycin B to the cells and keep in the selection media until all cells in the untransfected control condition die. Antibiotic selection concentration varies between cell types. We recommend plotting a titration vs. cell viability curve to determine the lowest concentration at which nearly all cells die within 7 – 12 days.

*If the cell number is low after 72 hours, start hygromycin B selection on the 5th day of transfection. This increases the chance of getting positive colonies. The addition of conditioned media (8:2 ratio of new media with filtered cultured media from the same type of cell) increases the survival of the clones.*

52 After all the cells in the control plate are dead, replace the media with regular conditioned media in all the transfected cells and control plates.

53 Colonies will appear in the plates after two or three weeks.



**Citation:** Kizhakke Mattada Sathyan, Thomas G. Scott, Michael J. Guertin (04/01/2020). Basic Protocol 2: Tagging a gene of interest with AID. https://dx.doi.org/10.17504/protocols.io.bdyei7te

54 Pick colonies similar to the steps 15 – 22 (from basic protocol 1) and screen for tagged clones by Western blotting.

*We use gene specific antibodies for detecting tagged proteins. The successful tagging results in an increase in protein size of 24 KD. Heterozygous integration will contain bands reflecting both the native and tagged proteins. The non-homologous end repair may produce read through into the 3'UTR or introduce stop codons before normal stop codon of the gene. This is observed as an unexpected size shift that is not ~24KDa. Confirm the integration and reading frame of the integrated AIDP2A-HygR or HygR-P2A-AID by PCR and sequencing.*

55 Select colonies from the Western blotting experiment and perform genomic DNA PCR using the same set of primers used for testing sgRNA efficiency according to steps 31 – 34, with the following change: PCR extension time should be increased to **00:04:00** to amplify the insert.

56 Electrophorese whole PCR product on [M]**1 %** agarose gel, excise the band using a clean scalpel and use the Qiagen gel purification kit to isolate genomic DNA (follow steps 2 – 12). Sequence the purified PCR product with the forward primer to confirm the reading frame.

*Successful integration of the AID-P2A-HygR or HygR-P2A-AID results in an addition of approximately 1785 bp to the PCR product. Heterozygous clones will have two bands, one with a length of genomic region between forward and reverse primer and the second with a length of genomic region between primers plus the 1785 bp. The homozygous integration will result in one band with a length of genomic region between primers plus the 1785 bp. Sequence the integrated DNA using the same forward primer, which will confirm successful integration.*





**Citation:** Kizhakke Mattada Sathyan, Thomas G. Scott, Michael J. Guertin (04/01/2020). Basic Protocol 2: Tagging a gene of interest with AID. https://dx.doi.org/10.17504/protocols.io.bdyei7te