



Published in final edited form as:

Adv Biosci Biotechnol. 2017 September ; 8(9): 311–323. doi:10.4236/abb.2017.89023.

TALEN-Mediated FLAG-Tagging of Endogenous Histone Methyltransferase DOT1L

Cheng An^{1,2,*}, Guangjing Zhu^{2,*}, Suzanne N. Martos², Xue Feng¹, Haimou Zhang³, Yankai Jia⁴, and Zhibin Wang^{2,3,5,6,#}

¹Guang'an Men Hospital, China Academy of Chinese Medical Sciences, Beijing, China

²Laboratory of Human Environmental Epigenomes, Department of Environmental Health Sciences, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA

³School of Life Sciences, Hubei University, Wuhan, China

⁴GENEWIZ Suzhou, Suzhou, China

⁵Fenxian Central Hospital, Shanghai, China

⁶Department of Oncology and Sidney Kimmel Comprehensive Cancer Center, School of Medicine, Johns Hopkins University, Baltimore, MD, USA

Abstract

Histone modification including H3 lysine 79 methylation (H3K79me) plays a key role during gene transcription and DNA damage repair. DOT1L, the sole methyltransferase for three states of H3K79me, is implicated in leukemia, co-lorectal cancer, and dilated cardiomyopathy. However, understanding of DOT1L and H3K79me in these pathways and disease pathogenesis has been limited due to the difficulty of working with DOT1L protein. For instance, locus-specific or genome-wide binding sites of DOT1L revealed by chromatin immunoprecipitation (ChIP)-based methods are necessary for inferring its functions, but high-quality ChIP-grade antibodies are currently not available. Herein we have developed a knock-in approach to tag endogenous DOT1L with 3 × Flag at its C-terminal domain to follow functional analyses. The knock-in was facilitated by using TALENs to induce a targeted double-strand break at the endogenous DOT1L to stimulate local homologous recombination at that site. The single cell colonies with successful knock-in were isolated and verified by different methods. We also demonstrated that tagged DOT1L maintains its normal function in terms of methylation and that the engineered cells would be very useful for further studies.

Keywords

DOT1L; Flag; TALEN; Knock-In

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <http://creativecommons.org/licenses/by/4.0/>

zwang47@jhu.edu.

*Both authors contributed equally to this project.

Conflict of Interest

Z.W serves as a consultant for CUSABIO (CusAb) Company, Wuhan, China.

1. Introduction

1.1. DOT1L and Its Functional Roles

Histone lysine methylations play a key role in gene transcription [1] [2] [3] and they are catalyzed by a group of histone lysine methyltransferases (KMTs). Yeast Dot1 (disruptor of telomeric silencing; or Kmt4) [4] or its mammalian homolog DOT1-Like (DOT1L) is the sole KMT in their respective genome with methylation activities toward H3K79 inside globular domain of histones. Both enzymes can catalyze mono-(me1), di-(me2), and trimethylation (me3) in a non-processive manner [5] [6]. In addition to histone methylation, DOT1L can methylate other protein factors for regulation. For example, methylation of androgen receptor at lysine 349 is involved in the regulation of androgen receptor responsive genes [7] [8]. DOT1L plays pivotal roles in embryonic development, hematopoiesis, and cardiac function [9]. Due to its role in development, it is not surprising to find that DOT1L is linked to many human diseases, including prostate cancer, colorectal cancer, and hypertension [7] [8] [10] [11] [12]. Interestingly, high level of DOT1L expression and H3K79me2 was a predictor of poor patient survival of colorectal cancer[11].

Considering the importance of DOT1L, the genome-wide location analysis (using ChIP-Seq) of DOT1L will provide detailed information to infer its function. However, the lack of specific, ChIP-Seq grade antibodies of DOT1L (Drs. Yi Zhang and T-p Chen, personal communications [13]) hinders such genome-wide analysis. To solve this lack of high quality antibody, one approach is to over express Flag- or Myc-tagged DOT1L and then use anti-Flag or Myc for ChIP assays. For example, we have used anti-Flag to map the location of DACH1, a protein with roles in breast cancers, in breast cancer cell line MDA- MB-231 with stably expressed Flag-DACH1 [14]. However, this over expression approach is limited with difficulty in maintaining a physiologically relevant expression level. Flag- or Myc-tagging the endogenous gene would avoid the aforementioned limitation. In addition, the small size of Flag sequences seem to seldom affect the target protein, and the availability of high affinity anti-Flag antibodies make the ChIP assays highly successful. Furthermore, endogenous tagging will reveal the true function of the protein compared to over expression studies[15] [16].

Current methods of knock-in are limited by low homologous recombination (HR) efficiency even when adeno-associated virus (AAV) was used as a vector to provide the donor plasmid for HR. Apart from the complexity of virus packaging, isolation of hundreds of colonies followed by verification by PCR is laborious and may result in only a few successful knock-in colonies. A targeted double strand break (DSB) has reported to increase HR efficiency by 2000 ~ 10,000 fold [16]. Herein, we induced a DSB using TALEN (described below) to facilitate tagging of endogenous DOT1L with 3 × Flag sequences to generate single colony cell lines that express DOT1L-3 × Flag endogenously.

1.2. TALEN-Mediated Genome Editing

Since the original reports on the creation and use of zinc finger nucleases (ZFNs) to stimulate locus-specific genome editing by Chandrasegaran group [17] [18], we have seen

dramatic progress in the field of genome engineering. It is laborious and time-consuming to generate highly specific ZFNs for desired target sites. The fortuitous discovery of transcriptional activator-like effectors (TALEs) [19] [20] has expanded and simplified the generation of custom TALE DNA-binding domains with programmable specificity [21] [22] [23]. Repeat monomers in TALE with amino acids at position 12 and 13 bind to a specific nucleotide (e.g., NI to A, HD to C, NG to T, and NN to G or A). The TALE motifs are then linked in tandem to generate custom TALE binding domains that target a specific sequence. These are then coupled to Fok I nuclease domain (to form TALENs) or other effector domains (such as transcription factors: TALE-TF) providing a versatile platform for achieving a wide variety of targeted genome manipulations specificity that include gene knock-out, knock-in, activate and repress gene expression, respectively [21] [22] [23] [24] [25]. A pair of TA-LENs separated by 14 – 20 bp are designed to induce a targeted DSB at the desired chromosomal locus, which will be repaired by one of the two cellular mechanisms: non-homologous end joining (NHEJ) or HR (when provided with an exogenous donor plasmid containing homologous sequences flanking the cut site).

Here, we report efficient tagging of the endogenous DOT1L with 3 × Flag using TALENs and a donor plasmid containing homologous sequences. We show successful knock-in of DOT1L-3 × Flag into HEK293 cells, which was verified by using different methods. However, low expression of endogenous DOT1L in HEK293 cells, limited the ChIP-Seq experiments for detecting genome-wide binding sites.

2. Materials and Methods

2.1. Construction of Donor Plasmid for Homologous Recombination

We followed a modified protocol from the method in reports [15] [16]. A ~1.1 kb genomic fragment upstream and 1.2 kb downstream of the stop codon of human DOT1L (transcript 006 of Ensembl release 69—October 2012, ENST00000440640) fragment was PCR amplified using platinum Taq polymerase High Fidelity (Invitrogen, USA) to form the left and right homologous arms, respectively. Refer to Table 1 for the information on primers. These were cloned into pAAV-USER-Neo-LoxP-3 × Flag vector (kind gift from Dr. Zhenghe Wang at Case Western Reserve University) using the USER system (NEW ENGLAND BioLabs, USA) as previously described elsewhere [15] [16]. This recombinant plasmid served as the DOT1L donor plasmid without packaging into AAV for the knock-in study. The donor plasmid was sequenced to verify the sequences of the left and right homology arms.

2.2. Construction of TALEN Pairs Targeting the Stop Codon Region of DOT1L

TALEN pairs targeting the stop codon regions of DOT1L were constructed based on the system developed in a previous report [22]. The plasmids for TALEN system was ordered from Addgene as denoted in the protocol [19]. We designed one pair of TALEN targeting the following sequence of the human DOT1L gene: 5-TATGCCGCGCACCTTTCCGG gggttaagccgcgataa AGACCTTGCTTAGCTAGCA-3 where the uppercase nucleotides show the binding sites for TALEN left (TN-1L, 5 end) and right (TN-1R, 3 end) arms, respectively; and the lowercase letters represent the spacer sequence between the left and

right arms of the TALEN pair. The stop codon (taa) is underlined. We also designed an additional TALEN pair (TN-2L and TN-2R) targeting 5-TGCCGCGCACCTTTTCGGGGggttaaagccgcgataaagaccttGCTT AGCTAGCAGTGCGTA-3. The combination of TN-1L and TN-2R generates a third TALEN pair for targeting. The genes encoding the TALE motifs are shown in Table 2.

2.3. Cell Culture and Transfection

HEK293 cells (ATCC[®] CRL-1573[™], US) were cultured in Dulbecco's Modified Eagle Medium containing 10% FBS and 1% penicillin and streptomycin (Life Technologies, USA), and maintained in a humidified incubator (5% CO₂) at 37°C.

For the knock-in (KI) study, HEK293 cells were transfected with 2 µg of TALEN left arm and right arm plasmids using GenJet[™] In Vitro DNA Transfection Reagent (Ver. II, SiganaGen[®] Laboratories). After transfection, the cells were cultured for 48 hours to extract genomic DNA to monitor the cutting efficiency of the TALEN pairs (TN-1L/1R; TN-2L/2R; TN-1L/TN-2R). For the knock-in study, HEK293 cells were transfected with 2 µg of TN2L and TN2R plasmids, and 2 µg donor plasmids using GenJet[™] In Vitro DNA Transfection Reagent.

2.4. Verification of TALEN Cutting Efficiency Using the Surveyor Nuclease

As reported previously[22], genomic DNA of TALEN pairs-transfected HEK293 cells were PCR amplified around the stop codon region with primer pair F2 and R2. The resulting PCR product contains a heterogeneous population (649 bp) of both TALEN-modified and unmodified DNA. Heteroduplexes of both amplified DNA were generated after denaturing and then re-annealing slowly. Surveyor nuclease (from Transgenomic, USA) cleaves the heteroduplexes at TALEN targeting sites (stop codon region) with mismatches to produce two fragments (~219 and 430 bp in size), while leaving homoduplexes intact. The two fragments were separated by gel electrophoresis and visualized under UV.

2.5. Monoclonal Knockin Cell Selection and Verification

After transfection with donor plasmid and TALEN plasmids, cells were selected with Geneticin 418 disulfate salt (G418, Singma-Aldrich, USA) for 48 hours, and then the cells were dispersed and serially-diluted in 96-well plates with medium containing G418 for single colony selection as reported elsewhere [15] [16]. The G418 resistant single cell clones were then screened for HR by PCR using the primer pairs F1/NR and NF/R1 for the left and right arms, respectively. Refer to Table 1 for the sequences of the primer pairs. Only cells that originated from a single cell colony containing the correct knock-in sequence were used in later study.

2.6. Western Blot

For Western blot, whole cell lysate of knock-in cell lines were extracted as described elsewhere [26]. 30 µg or 150 µg total protein extract was loaded on SDS-PAGE gel (BIO-RAD, USA) to separate the proteins according to their sizes. The proteins from the gel were then transferred to a PVDF membranes and incubated with primary antibodies of Flag (1:1000, F3165, Sigma), DOT1L (1:4000, A300-953A, Bethyl), H3K79me2 (1:1000,

Ab3594, Abcam), H3K79me3 (1:1000, Ab2621, Abcam) and β -actin (1:5000, A5441, Sigma) at 4 C overnight. After washing with $1 \times$ TBST (BIO-RAD, USA), the membranes were incubated with horseradish peroxidase-labeled secondary antibodies (GE healthcare, UK and EPITOMICS, USA) and then detected with ECL western blotting detection reagent (GE healthcare, UK).

3. Results

3.1. Overview of the Knock-In Strategy

Our objective was to tag C-terminal end of endogenous DOT1L with $3 \times$ Flag sequences and use anti-Flag antibody to pull down endogenous DOT1L for future analyses of DOT1L function in HEK293 cells. To achieve this, we first cloned ~ 1.1 kb sequences upstream and ~ 1.2 kb sequences downstream of the stop codon (including the stop codon) of DOT1L from HEK293 genomic DNA into pAAV-USER-Neo-LoxP- $3 \times$ Flag vector using the USER system to use as a donor plasmid (see Figure 1) [15] [16]. The cloned donor plasmid contains $3 \times$ Flag sequences just before the stop codon and neomycin resistance genes for selection of positive knock-in clones with G418 treatment. We used TALENs to induce a targeted DSB and transfected the donor plasmid into the cells to stimulate HR. As expected, the DSB was readily repaired by HR in presence of the donor plasmid; we observed highly efficient HR occurs at the DSB. Knock-in cells were further serially diluted and selected by G418 treatment. The correct knock-in single colonies were identified by PCR and sequencing. The strategy for the knock-in is shown in Figure 1.

3.2. TALEN Pair Design, Construction and Verification

Traditional knock-in methods rely on natural HR; however, the efficiency is very low even when using AAV as a vector for delivery of the donor plasmid. It has been reported that a targeted DSB could increase HR efficiency by 2000 \sim 10,000 fold [16]. We used the TALENs to create targeted DSB in the genome of HEK293 cells. We induced a TALEN-mediated targeted DSB at the stop codon region of DOT1L (Figure 1) within the genome of HEK293 cell line and provided the requisite donor plasmid for HR (incorporating a $3 \times$ Flag sequence after the stop codon of DOT1L gene). For knock-in strategy, we designed three pairs of TALENs (TN1L/1R, TN2L/2R, and TN1L/2R) targeting the stop codon region and verified their cleavage efficiency by surveyor nuclease assay (see Figure 2). All three TALEN pairs were shown to create a DSB, but with different efficiency: TN2L/2R pair was the most efficient among the three pairs and it was used for the knock-in study.

3.3. TALEN-Mediated Knock-In of $3 \times$ Flag Sequences to the Endogenous DOT1L

We transfected HEK293 cells with this TALEN pair (TN2L/2R) in presence of a donor plasmid (containing $3 \times$ Flag sequences before the stop codon of endogenous human DOT1L) for HR. Single colonies were selected with G418 selection. Two representative colonies were selected for further study. Two pairs of PCR primers (F1 and NR for left arm, NF and R1 for right arm) were used to verify the correct insertion of donor plasmid sequences within the genome (see Figure 1 & Figure 3). Only successful knock-in cell showed the right sized PCR products for both arms; the knock-in sequences of genomic were verified by Sanger sequencing (see Figure 4).

3.4. Verification of DOT1L-3 × Flag Expression after Knock-In

To verify the successful expression of DOT1L-3 × Flag, we did Western blot analysis with anti-Flag antibody using two representative knock-in single colonies. We were unable to detect DOT1L signal when we used 30 µg of total protein (a routine loading amount in our experiments) from whole cell lysate. We reasoned that the failed detection might be due to low expression level of DOT1L-3 × Flag. DOT1L-3 × Flag protein was then enriched by immunoprecipitation of whole cell lysates of knock-in cell lines with anti-Flag antibody. The Flag-enriched whole cell lysate was monitored by anti-DOT1L antibody; we successfully detected the DOT1L protein of the correct size in the knock-in cell lines, but not the control HEK293 cells (data not shown). Therefore, we reason that the expression of DOT1L-3 × Flag after knock-in is relatively low. When we increased the quantity of total proteins loaded into the gel to 150 µg, we could successfully detect the signal of DOT1L-3 × Flag using anti-Flag antibody cell lysates, without the need for enrichment (see Figure 5(a)).

3.5. Functional Verification: DOT1L-3 × Flag Knockin did not Change H3K79me2 and H3K79me3 Levels

DOT1L is the only enzyme that is responsible for (mono-, di- and tri-) methylation of H3K79. To verify that DOT1L-3 × Flag knock-in did not change the normal function of DOT1L, we extracted whole cell lysate of knock-in cell lines and examined H3K79me2 and H3K79me3 levels by Western blot. We found that knock-in did not change the expression of both methylation states of H3K79 (see Figure 5(b)), suggesting that DOT1L-3 × Flag functions similarly to endogenous DOT1L after KI.

4. Discussion

4.1. Pros and Cons of Knock-In

Although most systematic studies of mammalian proteins relied on ectopic expression of target proteins, these expressed proteins may lack important regulatory sequences such as 3' UTR or introns that are driven by constitutive or inducible promoters. Thus, they are quite different from the endogenous status and might provide results that are artifacts due to overexpression or non-endogenous regulation of tagged proteins. Epitope tagging of endogenous proteins circumvents this problem and has shown its power and convenience in different research areas such as epigenetics [12] [13] and proteomics [18]. However, the challenge is the relatively low targeting rates (2% ~ 5%) even when AAV virus was used for HR in cells [15] [16].

Endogenous epitope tagging provides an accurate platform for exploring various functions of proteins. However, most proteins have multiple transcripts which could be translated into proteins of different sizes that might function diversely. But, it is not uncommon that alternative transcripts share the same stop codon in the genome, which is tagged by 3 × Flag as in our case. Thus, one could tag a group of transcripts sharing the same stop codon. It is a fact that, to fully explore the profile of each transcript and its corresponding protein, laborious work needs to be done to tag different transcript groups with different stop codons. In this case, according to the Ensembl release 69 when the project was initiated, DOT1L has 9 splicing variants with 6 protein coding transcripts and among them, transcripts

DOT1L-006, -008 and -201 share the same stop codon, which was targeted for HR in our study. However, for the same gene, it is not always a good idea to choose the stop codon shared by most transcripts since they are not necessarily similarly abundant in mRNA level or functionally importantly at the protein level. RNA-Seq or microarray data would be helpful when it is available for the target genes. Our results here only could be explained by the 3 transcripts mentioned above and the lower expression of these 3 transcripts may be one reason for the failed ChIP-Seq experiment. Our ability to detect peaks could be limited by insufficient reads to detect a true difference in this model. Low endogenous expression of DOT1L combined with the non-specific binding of anti-Flag antibody could explain our inability to identify peaks.

4.2. Genomic Editing Techniques

The discovery of Zinc-finger nuclease [17] ushered in the era of genome engineering also known as genome editing, which was expanded further by TALENs [22]. The development of CRISPR/Cas system has made genome editing much simpler and easier [23]. In this study, we improved the previous knock-in technique [15] [16] by using TALENs, which creates double strand breakages (DSBs) of the targeted gene and was shown to increase the recombination efficiency significantly in many genomic regions [26]. We successfully tagged DOT1L with 3 × Flag with the USER system and further verified its proper function in this study. However, the lower expression of endogenous DOT1L in HEK293 cells limited our ChIP-Seq experiments with anti-Flag. A different cell line with highly expressed DOT1L will be a better model for future studies.

4.3. Key Steps for This Strategy

High efficiency knock-in is ascribed to the targeted DOT1L DSB using the TA-LENs. We suggest several improvements for efficient knock-in. First, the high sequence similarity between TALE motifs hinders rapid construction of the TALEN pairs. We purified TALE hexamers away from monomers after PCR reaction using gel extraction and PCR product purification kits. TALENs generated using these TALE hexamers almost always yield the correct genes, which has been confirmed by sequencing. Second, seed cells are no more than 60% confluence before transfection using TALEN pair and donor plasmid. Our suggestion is direct conversion of normal medium to that containing 1 mg/ml G418 after 48 hours. It is important that avoid treatment with trypsin. Third, pick colony with tip to resuspend in medium containing 1 mg/ml G418 and then serially dilute in a 96 well plate. After 12 hours, we need to monitor the wells containing single cell colonies under an inverted microscope to guarantee the positive cells arose from a single cell colony.

5. Conclusion

The strategy for knock-in based on the TALEN technique we showed is a high efficiency and successful. It is facilitated to induce a targeted double-strand break at the endogenous DOT1L to stimulate local homologous recombination at that site. The single cell colony with tagged DOT1L maintains its normal function in terms of methylation and that the engineered cells would be very useful for further studies.

Acknowledgments

We thank Dr. Zhenhe Wang and Peng Zhang in his group at Case Western Reserve University for providing donor plasmid backbones and all their technique support. TALEN system was greatly supported by Le Cong of Dr. Feng Zhang's group in Harvard Medical School. We also thank our colleague Dr. Srinivasan Chandrasegaran for his kind suggestions and help. This work was partially supported by NIEHS R01 ES025761 and NIEHS U01 ES026721 in the Wang laboratory. C.A. is supported by GAMH2014S301. S.M. is supported by NIEHS T32 ES07141.

References

1. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, et al. High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell*. 2007; 129:823–837. <https://doi.org/10.1016/j.cell.2007.05.009>. [PubMed: 17512414]
2. Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, et al. Combinatorial Patterns of Histone Acetylations and Methylations in the Human Genome. *Nature Genetics*. 2008; 40:897–903. <https://doi.org/10.1038/ng.154>. [PubMed: 18552846]
3. Wang Z, Schones DE, Zhao K. Characterization of Human Epigenomes. *Current Opinion in Genetics & Development*. 2009; 19:127–134. <https://doi.org/10.1016/j.gde.2009.02.001>. [PubMed: 19299119]
4. Singer MS, Kahana A, Wolf AJ, Meisinger LL, Peterson SE, Goggin C, et al. Identification of High-Copy Disruptors of Telomeric Silencing in *Saccharomyces cerevisiae*. *Genetics*. 1998; 150:613–632. [PubMed: 9755194]
5. Frederiks F, Tzouros M, Oudgenoeg G, van Welsem T, Fornerod M, Krijgsveld J, et al. Nonprocessive Methylation by Dot1 Leads to Functional Redundancy of Histone H3K79 Methylation States. *Nature Structural & Molecular Biology*. 2008; 15:550–557. <https://doi.org/10.1038/nsmb.1432>.
6. Shanower GA, Muller M, Blanton JL, Honti V, Gyurkovics H, Schedl P. Characterization of the Grappa Gene, the Drosophila Histone H3 Lysine 79 Methyltransferase. *Genetics*. 2005; 169:173–184. <https://doi.org/10.1534/genetics.104.033191>. [PubMed: 15371351]
7. Chung S, Nakagawa H, Uemura M, Piao L, Ashikawa K, Hosono N, et al. Association of a Novel Long Non-Coding RNA in 8q24 with Prostate Cancer Susceptibility. *Cancer Science*. 2011; 102:245–252. <https://doi.org/10.1111/j.1349-7006.2010.01737.x>. [PubMed: 20874843]
8. Petrovics G, Zhang W, Makarem M, Street JP, Connelly R, Sun L, et al. Elevated Expression of PCGEM1, a Prostate-Specific Gene with Cell Growth-Promoting Function, Is Associated with High-Risk Prostate Cancer Patients. *Oncogene*. 2004; 23:605–611. <https://doi.org/10.1038/sj.onc.1207069>. [PubMed: 14724589]
9. Nguyen AT, Zhang Y. The Diverse Functions of Dot1 and H3K79 Methylation. *Genes & Development*. 2011; 25:1345–1358. <https://doi.org/10.1101/gad.2057811>. [PubMed: 21724828]
10. Kryczek I, Lin Y, Nagarsheth N, Peng D, Zhao L, Zhao E, et al. IL-22(+)CD4(+) T Cells Promote Colorectal Cancer Stemness via STAT3 Transcription Factor Activation and Induction of the Methyltransferase DOT1L. *Immunity*. 2014; 40:772–784. <https://doi.org/10.1016/j.immuni.2014.03.010>. [PubMed: 24816405]
11. Duarte JD, Zineh I, Burkley B, Gong Y, Langae TY, Turner ST, et al. Effects of Genetic Variation in H3K79 Methylation Regulatory Genes on Clinical Blood Pressure and Blood Pressure Response to Hydrochlorothiazide. *Journal of Translational Medicine*. 2012; 10:56. <https://doi.org/10.1186/1479-5876-10-56>. [PubMed: 22440088]
12. Zhang X, Guo C, Chen Y, Shulha HP, Schnetz MP, LaFramboise T, et al. Epitope Tagging of Endogenous Proteins for Genome-Wide ChIP-Chip Studies. *Nature Methods*. 2008; 5:163–165. <https://doi.org/10.1038/nmeth1170>. [PubMed: 18176569]
13. Jones B, Su H, Bhat A, Lei H, Bajko J, Hevi S, et al. The Histone H3K79 Methyltransferase Dot1L Is Essential for Mammalian Development and Heterochromatin Structure. *PLoS Genetics*. 2008; 4:e1000190. <https://doi.org/10.1371/journal.pgen.1000190>. [PubMed: 18787701]
14. Zhou J, Wang C, Wang Z, Dampier W, Wu K, Casimiro MC, Chepelev I, Popov VM, Quong A, Tozeren A, Zhao K, Lisanti MP, Pestell RG. Attenuation of Forkhead Signaling by the Retinal Determination Factor DACH1. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103:1153–1158. <https://doi.org/10.1073/pnas.0510111103>. [PubMed: 16441111]

- States of America. 2010; 107:6864–6869. <https://doi.org/10.1073/pnas.1002746107>. [PubMed: 20351289]
15. Wang Z. Epitope Tagging of Endogenous Proteins for Genome-Wide Chromatin Immunoprecipitation Analysis. *Methods Mol Biol.* 2009; 567:87–98. https://doi.org/10.1007/978-1-60327-414-2_6. [PubMed: 19588087]
 16. Taghian DG, Nickoloff JA. Chromosomal Double-Strand Breaks Induce Gene Conversion at High Frequency in Mammalian Cells. *Mol Cell Biol.* 1997; 17:6386–6393. <https://doi.org/10.1128/MCB.17.11.6386>. [PubMed: 9343400]
 17. Kim YG, Cha J, Chandrasegaran S. Hybrid Restriction Enzymes: Zinc Finger Fusions to Fok I Cleavage Domain. *Proceedings of the National Academy of Sciences of the United States of America.* 1996; 93:1156–1160. <https://doi.org/10.1073/pnas.93.3.1156>. [PubMed: 8577732]
 18. Bibikova M, Carroll D, Segal DJ, Trautman JK, Smith J, Kim YG, Chandrasegaran S. Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases. *Molecular and Cellular Biology.* 2001; 21:289–297. <https://doi.org/10.1128/MCB.21.1.289-297.2001>. [PubMed: 11113203]
 19. Ousterout DG, Gersbach CA. The Development of TALE Nucleases for Biotechnology. *Methods in Molecular Biology.* 2016; 1338:27–42. https://doi.org/10.1007/978-1-4939-2932-0_3. [PubMed: 26443211]
 20. Ul Ain Q, Chung JY, Kim YH. Current and Future Delivery Systems for Engineered Nucleases: ZFN, TALEN and RGEN. *Journal of Controlled Release.* 2015; 205:120–127. <https://doi.org/10.1016/j.jconrel.2014.12.036>. [PubMed: 25553825]
 21. The Runners-Up: Genome Cruise Missile. *Science.* 2012; 338:1525–1532. <https://doi.org/10.1126/science.338.6114.1525>. [PubMed: 23258865]
 22. Sanjana NE, Cong L, Zhou Y, Cunniff MM, Feng G, Zhang F. A Transcription Activator-Like Effector Toolbox for Genome Engineering. *Nature Protocols.* 2012; 7:171–192. <https://doi.org/10.1038/nprot.2011.431>. [PubMed: 22222791]
 23. Cong L, Zhou R, Kuo YC, Cunniff M, Zhang F. Comprehensive Interrogation of Natural TALE DNA-Binding Modules and Transcriptional Repressor Domains. *Nature Communications.* 2012; 3:968. <https://doi.org/10.1038/ncomms1962>.
 24. Kim SK, Jung I, Lee H, Kang K, Kim M, Jeong K, et al. Human Histone H3K79 Methyltransferase DOT1L Methyltransferase Binds Actively Transcribing RNA Polymerase II to Regulate Gene Expression. *The Journal of Biological Chemistry.* 2012; 287:39698–39709. <https://doi.org/10.1074/jbc.M112.384057>. [PubMed: 23012353]
 25. Song J, Hao Y, Du Z, Wang Z, Ewing RM. Identifying Novel Protein Complexes in Cancer Cells Using Epitope-Tagging of Endogenous Human Genes and Affinity-Purification Mass Spectrometry. *Journal of Proteome Research.* 2012; 11:5630–5641. [PubMed: 23106643]
 26. Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, et al. Genetic Engineering of Human Pluripotent Cells Using TALE Nucleases. *Nature Biotechnology.* 2011; 29:731–734. <https://doi.org/10.1038/nbt.1927>.

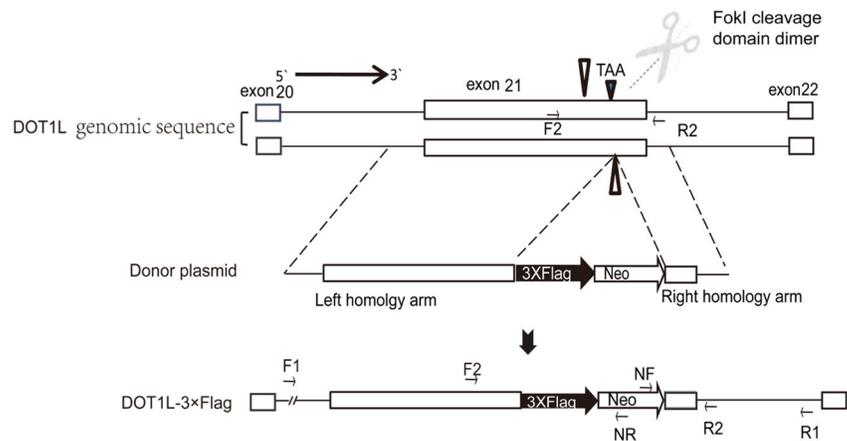


Figure 1.

Schematic diagram of knockin strategy. The strategy involved 3 steps: first, TALEN pairs were used to create DSB at the stop codon region of DOT1L to facilitate HR. The left and right arm of TALEN pairs recognize regions before and after the stop codon, and dimerized FokI endonuclease generates a site specific double strand break to facilitate DOT1L-3 × Flag gene knock-in through HR. Second, donor plasmid contains a left and right arm homologous to sequences in human DOT1L, flanking a Neo-Lox P-3 × Flag cassette was constructed. Donor plasmid was transfected along with the TALEN pair plasmids to induce a targeted DSB for HR. Third, successful knock-in single colony was picked and verified by PCR. Note that drawing is for viewing purpose only and it is not proportional to the size of genomic sequences.

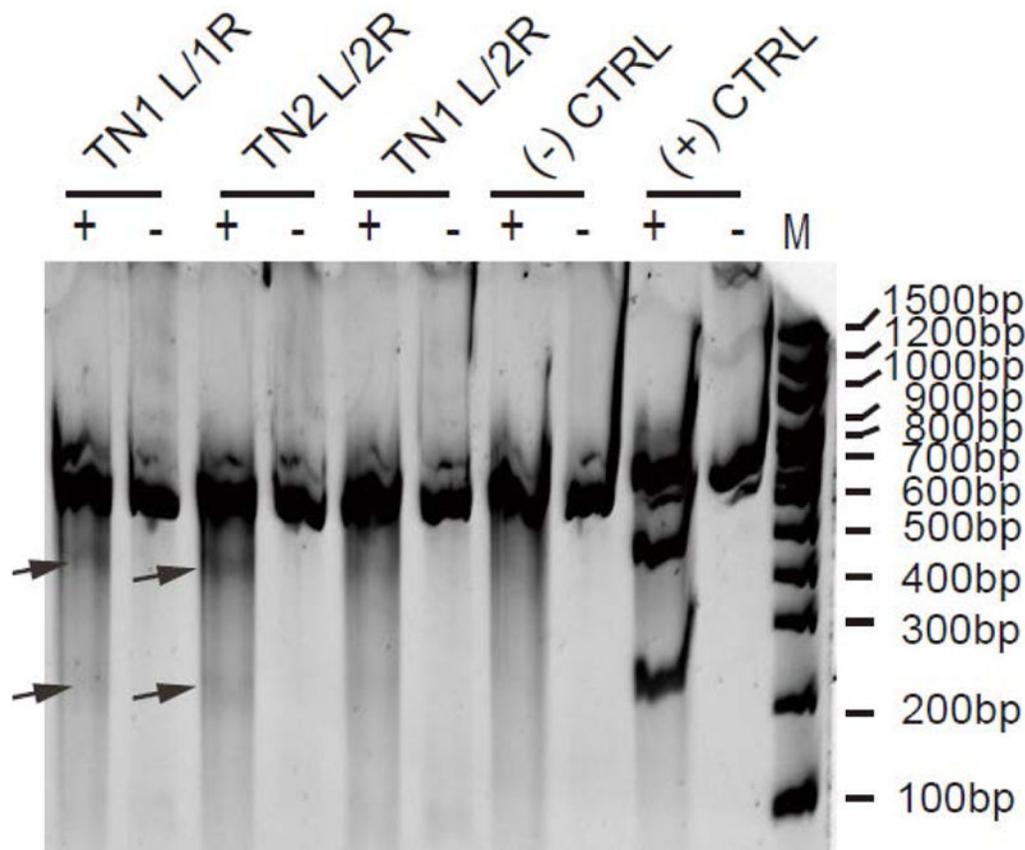


Figure 2.

Surveyor nuclease assay to verify the DSB induced by TALEN pairs. Genomic DNA of HEK293 cells transfected with TALEN pair plasmids were PCR amplified around the stop codon regions using primer pair F2 and R2 as shown in Figure 1. Surveyor nuclease treatment of successfully KO cells produces 2 segments (~219 bp & 430 bp, arrows). “+” and “-” represent with or without surveyor nuclease treatment, respectively. Negative and positive controls in the kit were verified the same way, with or without surveyor nuclease treatment. Surveyor nuclease digestion products were 217 bp and 416 bp for the positive control.

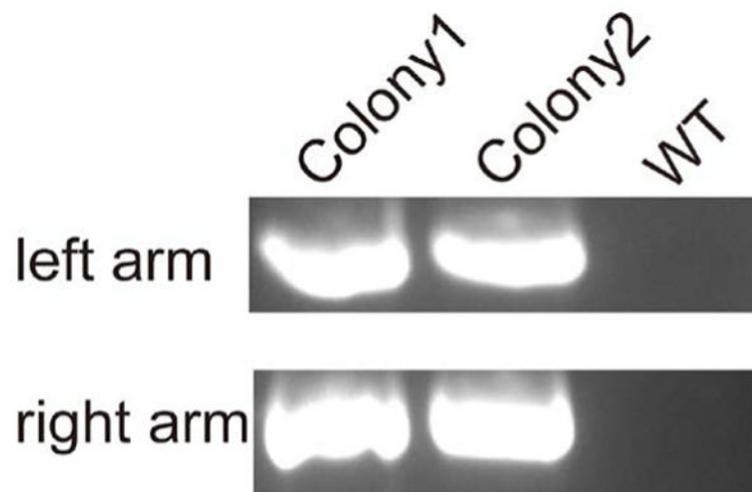


Figure 3. PCR verification of successful knock-in single colony cell lines. Genomic DNAs of single colony cells after knock-in were used to screen with primer pairs F1/NR and NF/R1, producing expected PCR products of 1712 bp and 1750 bp, respectively, confirming the correct HR at the DOT1L site.

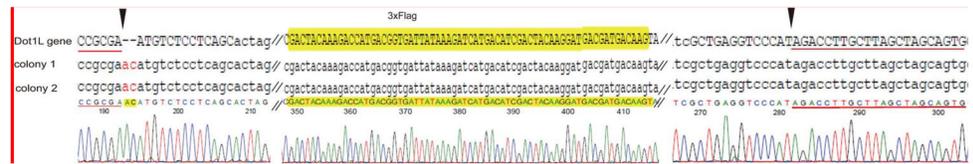


Figure 4.

Sequencing results from successfully KI cell. Left panel shows the junction of right flank of endogenous and left flank of plasmid and “ac” were added to guarantee fusion gene in frame. Red underline represents the sequence of Dot1L-P1R. Middle panel shows the sequence of 3 × Flag knocked in. Right panel shows the junction of left flank of endogenous and right flank of plasmid. Red underline represents the sequence of Dot1L-P2F.

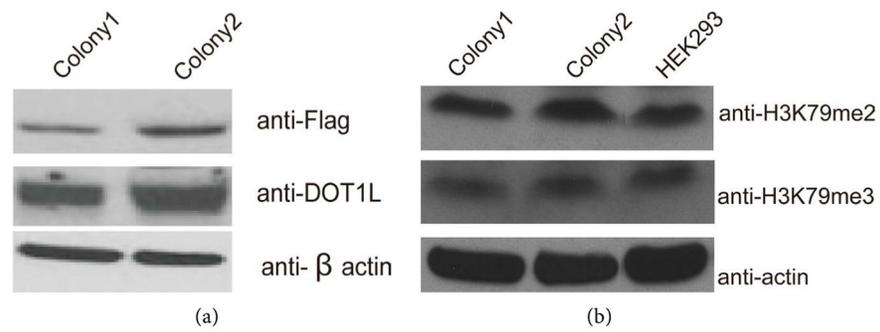


Figure 5. Function verification of DOT1L-3 × Flag knock-in cell lines (a). Whole cell lysate of single colony cell lines expressing DOT1L-3 × Flag were used for Western blot using anti-Flag antibodies (150 μg whole cell lysate). β-actin was used as a control (bottom). (b). Western blot of H3K79me2 and H3K79me3 were done using whole cell lysate of DOT1L-3 × Flag knockin cell lines and the knock-in did not change the methylation status of H3K79me2 and H3K79me3.

Table 1

Primers sequence for knock in and verifying test.

Primer Name	Sequences	Size (bp)	Uses
DOT1L-P1R	GGA GAC A/ideoxyU/GTTCGCGGCTTAACCCCGAAA	1079	Left arm
DOT1L-P1F	GGG AAA G/ideoxyU/TGCACGGTGGCGAACTCCAG		
DOT1L-P2R	GGC ATA G/ideoxyU/TCCTAAGAGACCCAGCATAG	1152	Right arm
DOT1L-P2F	GGTCCCA/ideoxyU/AGACCTTGCTTAGCTAGCAG		
F1	CAGGTTCCCTTCCGACTCT	1713	picking positive colony
NR	GTTGTGCCCAGTCATAGCCG		
NF	TCTGGATTCATCGACTGTGG	1750	picking positive colony
R1	TAGTTACCTGCAGAAGGGCA		
F2	CCGCCTGCTAACGCCTCTTT	649	Surveyor nuclease cutting
R2	ACGCCACCCGTCATGAGTGA		

Table 2

TALEN pairs information.

TN pairs	Arms/Sequences	Sequences
TN1	Left	NI NG NH HD HD NH HD NH HD NI HD HD NG NGNG HD NH NH
	Right	NH HD NG NI NH HD NG NI NI NH HD NI NI NH NH NG HD NG
	Sequence	T ATGCCGCGCACCTTTCGG gggtaagccgcgataa AGACCTTGCTTAGCTAGC A
TN2	Left	NH HD HD NH HD NH HD NI HD HD NG NGNG HD NH NHNH
	Right	NI HD NH HD NI HD NG NH HD NG NI NH HD NG NI NI NH HD
	Sequence	T GCCGCGCACCTTTCGGGG gttaagccgcgataaagacctt GCTTAGCTAGCAGTGC GT A

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript