HHMI Author Manuscript

TALEN-mediated editing of the Mouse Y Chromosome

Haoyi Wang[1,7], **Yueh-Chiang Hu**[1,2,7], **Styliani Markoulaki**[1], **G. Grant Welstead**[1], **Albert W. Cheng**[1,4], **Chikdu S. Shivalila**[1,3], **Tatyana Pyntikova**[1,2], **Daniel B. Dadon**[1,3], **Daniel F. Voytas**[5], **Adam J. Bogdanove**[6], **David C. Page**[1,2,3,8], and **Rudolf Jaenisch**[1,3,8]

[1] Whitehead Institute for Biomedical Research, Cambridge, Massachusetts, USA

[2] Howard Hughes Medical Institute

[3] Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

[4] Computational and Systems Biology Program, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

[5] Department of Genetics, Cell Biology and Development, University of Minnesota, Saint Paul, Minnesota, USA

[6] Department of Plant Pathology and Plant-Microbe Biology, Cornell University, Ithaca, New York, USA

Abstract

The functional study of Y chromosome genes has been hindered by a lack of mouse models with specific Y chromosome mutations. We used transcription activator-like effector nuclease (TALEN)-mediated gene editing in mouse embryonic stem cells (mESCs) to produce mice with targeted gene disruptions and insertions in two Y-linked genes – *Sry* and *Uty*. TALEN-mediated gene editing is a useful tool for dissecting the biology of the Y chromosome.

The Y chromosome has unique structure and gene content and has evolved into a chromosome that is highly specialized for male sex differentiation and fertility[1, 2]. A comprehensive approach to study the function of Y-linked genes in mice is needed to unravel the biology of the Y chromosome. Perhaps due to the unique structural features of the Y chromosome, conventional gene targeting strategies in mESCs to generate mutations in Y-linked genes have been unsuccessful. Therefore, our understanding of the functions of murine Y-linked genes is limited to insights gained from studies of mice that carry spontaneous deletions, random gene trap insertions or autosomal transgenes[3-5]. Although two Y-linked knockout mESC clones and one transgenic mouse line have been generated using an insertional targeting strategy (a method causing DNA duplications around the targeting site)[6, 7], no mouse with targeted gene knockout and knock-in on the Y chromosome has been reported. Here we report using TALEN-mediated gene editing[8, 9] to

efficiently manipulate genes on the mouse Y chromosome and produce mice with mutations in two different Y-linked genes.

We first engineered TALENs (Fig. 1A, Supplementary Table 1) directed against the mouse *Sry* gene. *Sry* has a well-defined function in testis determination and is therefore an appropriate target for proof-of-principle experiments. Two pairs of TALENs were generated to target the high mobility group (HMG) DNA binding domain of *Sry*. We used the Surveyor assay[10] to test the nuclease activity of these TALENs on the *Sry* gene in V6.5 XY mESCs. TALEN pairs 1 and 2 showed gene modification efficiencies of 15% and 20%, respectively (Supplementary Fig. 1). Following TALEN transfection, mESCs were plated at low density without selection, and clones were picked and genotyped by Southern blot analysis. Because *Sry* TALEN pair 2 cleaves a region that includes a BsaJI site, successfully targeted clones could be identified by a loss of the BsaJI site, using Southern blot analysis (Fig. 1A and B). In three independent targeting experiments we screened 200 mESC clones and obtained five targeted clones. Each clone harbored a single *Sry* gene deletion ranging from 11 to 540 bp (Fig. 1A). Four of the five targeted mESC clones retained a normal chromosome configuration (Supplementary Fig. 2).

To generate mice carrying a targeted mutation (tm) of the *Sry* gene, *Sry^{tm1}* (41-bp deletion), *Sry^{tm2}* (11-bp deletion), and *Sry^{tm4}* (540-bp deletion) mESCs were injected into tetraploid blastocysts, which were transferred to pseudopregnant females. Embryos were examined at E14.5 (12 days after transfer) and pups were delivered by Caesarian section (17 days after transfer). As expected, all E14.5 embryos (n=7) and full-term pups (n=17) that we derived from these three mESC clones displayed sex-reversal such that chromosomal (XY) males were anatomically female (Fig. 1C and D). We did not observe any evidence of testicular differentiation in either embryos or pups. Adult *Sry*-targeted mice (anatomic females) showed reduced fertility, but transmitted the *Sry*-mutated Y chromosome to offspring (Fig. 1D and Supplementary Table 2). These data confirm a previous report of anatomic sex reversal in *Sry^{dl1Rlb}* mice carrying a spontaneous 11-kb deletion encompassing the *Sry* locus[3]. To confirm that sex reversal was caused by the *Sry* mutation and not other genetic changes, we crossed *Sry^{tm1}* females with *Sry^{dl1Rlb}*;Tg(*Sry*)2Ei males, a previously described *Sry* transgenic line[11], to produce *Sry^{tm1}*;Tg(*Sry*)2Ei progeny. Consistent with the previous report[12], pups with the *Sry* mutation that also carried the *Sry* transgene (*Sry^{tm1}*;Tg(*Sry*)2Ei) developed as anatomic males (Fig. 1D and Supplementary Table 2), supporting the conclusion that the anatomic sex reversal observed in *Sry^{tm1}* mice was due to the absence of *Sry*.

To test for potential non-specific mutations induced by introduction of the TALENs, we predicted the top 15 potential off-targets of TALEN pair 2 (Supplementary Table 3) using the position weight matrix (PWM) of natural TAL effectors[13] (Materials and Methods). We PCR-amplified and sequenced these 15 loci in *Sry^{tm1}*, *Sry^{tm2}*, and *Sry^{tm4}* mESC clones and found no mutations.

To further establish the utility of TALEN-mediated targeting for the Y chromosome we produced a GFP knock-in allele at the *Sry* locus. We introduced TALEN pair 2, along with a donor construct that contained a promoterless GFP and a puromycin selection marker flanked by short homologous arms (700 and 385 bp for 5' and 3' arms, respectively), into V6.5 mESCs (Fig. 1E). After transfection, mESCs were selected with puromycin and gancyclovir. In three independent experiments we screened 300 mESC clones and obtained three correctly targeted clones, all of which were confirmed by Southern blotting, PCR genotyping (Fig. 1F and Supplementary Fig. 3) and sequencing (data not shown). Since the Sry promoter drives GFP, we tested whether its expression recapitulates endogenous *Sry* expression. We performed semi-quantitative RT-PCR to detect GFP transcript levels in

various embryonic tissues. We dissected E12.0 embryos generated by tetraploid complementation and found that they were all anatomically female, indicating *Sry* gene inactivation. We then confirmed that GFP was expressed in the gonads and brain, but not in other tissues (Fig 1G), consistent with previous reports[12, 14]. The generation of *Sry*-GFP knock-in mice demonstrates the feasibility of insertion of a transgene into a specific locus on the Y chromosome.

Successful manipulation of the *Sry* gene prompted us to test whether the same targeting strategies could be applied to other Y-linked genes with similar efficiency. We chose *Uty* as the second target, because unlike *Sry, Uty* is expressed in mESCs[5], allowing to assay the effect of targeting. We designed two pairs of TALENs that target the first exon of *Uty* (Fig. 2A and Supplementary Table 1). Although the TALEN pair used to target *Uty* had a low efficiency of cutting when tested in a Surveyer assay, we were able to isolate one correctly targeted clone (*Uty^{tm1}* that contains a 167-bp insertion derived from the TALEN expression plasmid) out of 192 mESC clones picked. To improve the likelihood that the tested mESC clones had received the TALEN pair, we co-transfected *Uty* TALEN pair 2 along with a puroR-containing plasmid into mESCs, and selected for cells with transient puromycin resistance for two days. Drug selection dramatically increased the targeting efficiency from ~0.5% to ~9% (five of 56 clones were targeted), as confirmed by Southern blot analysis (Fig. 2B). We then derived mice from *Uty*[tm1] mESCs through tetraploid complementation. The *Uty*[tm1] mice were viable and fertile. When mated with *Utx* heterozygous mutant females[15], no *Utx^{tm}/Uty^{tm1}* double mutant offspring were obtained, while *Utx* and *Uty* single hemizygous-mutant males were born (Supplementary Table 4). These results suggest that *Utx* and *Uty* function redundantly in embryonic development, which is consistent with the previously reported phenotypes of gene trap-derived *Utx* and *Uty* mutant animals[5].

Using a similar strategy we generated an *Uty*-GFP knock-in allele by introducing *Uty* TALEN pair 2 and a donor construct containing a promoterless GFP and a puromycin selection marker flanked by short homologous arms (832 and 889 bp for 5' and 3' arms, respectively) into V6.5 mESCs (Fig. 2C). From three independent experiments, we screened 300 mESC clones and obtained six correctly targeted clones in total (~2% targeting efficiency), all confirmed by Southern blotting, PCR genotyping (Fig. 2D and Supplementary Figs. 4 and 5), and sequencing (data not shown). Because GFP is fused in-frame with the first 13 codons of *Uty*, all correctly targeted mESC clones were GFP positive (Fig 2E and F) as expected. We then generated animals from these *Uty*-GFP mESCs through tetraploid complementation. We found that GFP was undetectable by eye in embryos and newborn pups, which is consistent with a previous report that Uty is expressed at a low level in mouse embryos[5]. However, we were able to detect low-level GFP expression in mouse embryonic fibroblasts (MEFs) derived from the E13.5 *Uty*-GFP embryos using flow cytometry (Fig. 2F). This was confirmed by quantitative RT-PCR analysis, which detected a much lower expression level of the *Uty*-GFP transcript in MEFs compared to mESCs (Fig. 2G). To determine whether this differential expression is characteristic of the wild-type *Uty* locus, we compared *Uty* transcript levels in wild-type MEFs and mESCs and found a similar expression pattern (Fig. 2G).

We demonstrate that both *Sry* and *Uty* can be efficiently targeted by TALEN-mediated gene editing strategies, enabling the generation of mice carrying Y-linked gene mutations. The system described here provides a novel and general approach for genetic manipulation of the Y chromosome, which has not been possible with conventional gene targeting approaches. Thus, TALEN-mediated gene editing will allow the study of Y-chromosome biology by genetic manipulation in mice and other species.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Skaletsky H, et al. Nature. 2003; 423:825–837. [PubMed: 12815422]

2. Hughes JF, et al. Nature. 2012; 483:82–86. [PubMed: 22367542]

3. Lovell-Badge R, Robertson E. Development. 1990; 109:635–646. [PubMed: 2401216]

4. Mazeyrat S, et al. Nat Genet. 2001; 29:49–53. [PubMed: 11528390]

5. Shpargel KB, Sengoku T, Yokoyama S, Magnuson T. PLoS Genet. 2012; 8:e1002964. [PubMed: 23028370]

6. Rohozinski J, Agoulnik AI, Boettger-Tong HL, Bishop CE. Genesis. 2002; 32:1–7. [PubMed: 11835668]

7. Pirottin D, et al. Proc Natl Acad Sci U S A. 2005; 102:6413–6418. [PubMed: 15851664]

8. Miller JC, et al. Nat Biotechnol. 2011; 29:143–148. [PubMed: 21179091]

9. Cermak T, et al. Nucleic Acids Res. 2011; 39:e82. [PubMed: 21493687]

10. Guschin DY, et al. Methods Mol Biol. 2010; 649:247–256. [PubMed: 20680839]

11. Washburn LL, Albrecht KH, Eicher EM. Genetics. 2001; 158:1675–1681. [PubMed: 11514455]

12. Koopman P, Gubbay J, Vivian N, Goodfellow P, Lovell-Badge R. Nature. 1991; 351:117–121. [PubMed: 2030730]

13. Moscou MJ, Bogdanove AJ. Science. 2009; 326:1501. [PubMed: 19933106]

14. Mayer A, Mosler G, Just W, Pilgrim C, Reisert I. Neurogenetics. 2000; 3:25–30. [PubMed: 11085593]

15. Welstead GG, et al. Proc Natl Acad Sci U S A. 2012; 109:13004–13009. [PubMed: 22826230]
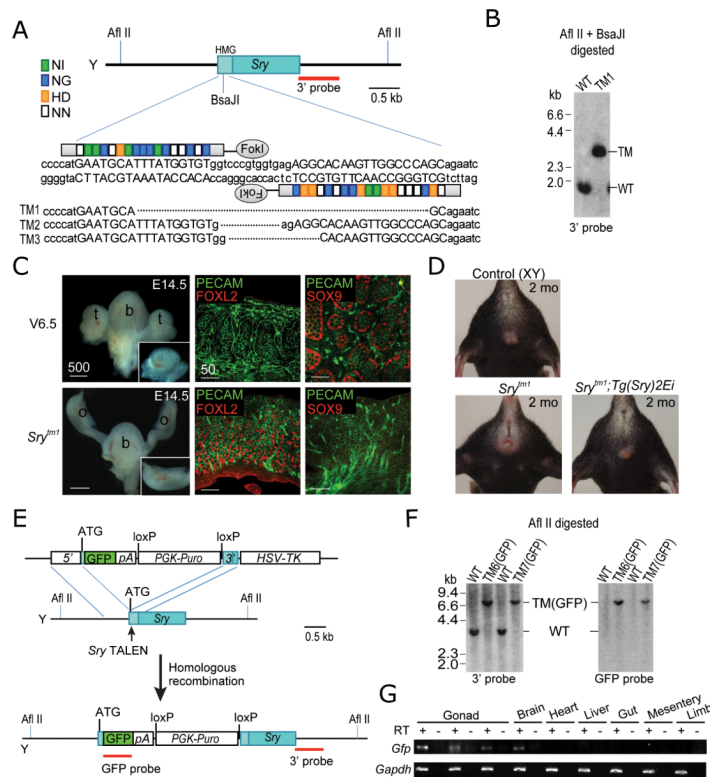
Figure 1. Genetic modification of *Sry* using TALENs

(A) Schematic of *Sry* TALEN pair 2 and its recognition sequence in the high mobility group (HMG) domain of *Sry*. TAL repeats are color-coded to represent each of four repeat variable di-residues (RVDs); each RVD recognizes one corresponding DNA base (NI = A, NG = T, HD = C, NN = G). Nucleotides bound by TALENs are capitalized. Shown below are clones (targeted mutation [TM] alleles 1-3) with *Sry* deletions induced by TALENs indicated using dashed lines. *Sry^{tm4}* (540-bp deletion) and *Sry^{tm5}* (440-bp deletion) clones are not shown. (B) Southern analysis of targeted alleles. AflII/BsaJI-digested genomic DNA was hybridized with a 3' probe. Expected fragment size: WT = 1.7 kb WT, TM = 3.2 kb. (C) E14.5 control embryo derived from parental V6.5 mESCs developed as ananatomic male, with testes (t), whereas *Sry^{tm1}* embryo developed as an anatomic female, with uterus and ovaries (o). b, bladder. Section of control gonad shows Sertoli cell marker expression (SOX9) and testicular cord formation, while section of *Sry^{tm1}* gonad shows granulosa cell marker expression (FOXL2). PECAM marks both endothelial and germ cells. Scale bar unit: · m. (D) *Sry^{tm1}*-bearing offspring of *Sry^{tm1}* females also exhibited female external genitalia and mammary glands (lower left). Anatomic sex reversal of *Sry^{tm1}* mice was rescued by the *Sry* transgene (lower right). (E) Schematic overview of strategy to generate *Sry*-GFP knock-in alleles. (F) Southern analysis of knock-in alleles. Afl II-digested genomic DNA was hybridized with 3' probe or internal GFP probe. Expected fragment size: WT = 3.9 kb, TM(GFP) = 7.0 kb. (G) RT-PCR analysis of GFP transcript expression in tissues from E12.0 *Sry*-GFP embryos. Internal controls without RT confirmed the absence of genomic DNA contamination. *Gapdh* was used as reference.
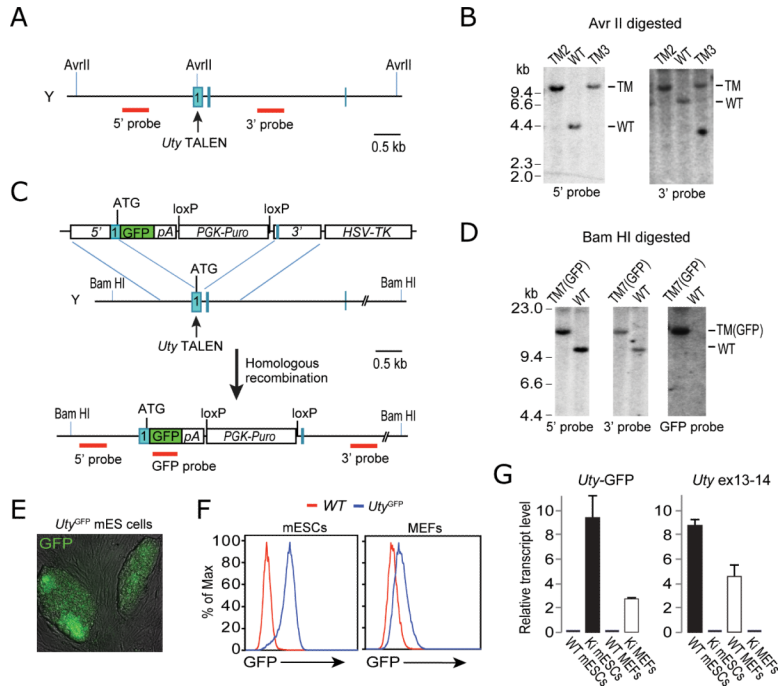
Figure 2. Genetic modification of *Uty* using TALENs
(A) Schematic overview of targeting strategy. (B) Southern analysis of AvrII-digested genomic DNA with 5' probe or 3' probe. Expected fragment size: WT , 4.7 kb (5' probe) and 7.7 kb (3' probe); TM = 12.5 kb for both probes. (C) Schematic overview of strategy to generate *Uty*-GFP knock-in alleles. (D) Southern analysis of BamHI-digested genomic DNA with 5' probe, 3' probe or internal GFP probe. Expected fragment size: WT = 10.4 kb, TM(GFP) = 13.5 kb. (E) Bright field/GFP overlay image shows that *Uty*-GFP mESCs expressed fluorescent protein. (F) FACS analysis for GFP fluorescence in *Uty*-GFP mESCs and MEFs, compared with wild-type counterparts. (G) Quantitative analysis of *Uty*-GFP and *Uty* transcript levels in cells by RT-PCR using primers spanning exon 1-GFP (left) and exons 13-14 (right), respectively. Ki = knock-in.