HHS Public Access

Author manuscript

Dev Biol. Author manuscript; available in PMC 2024 December 10.

Published in final edited form as:

Dev Biol. 2023 December ; 504: 120–127. doi:10.1016/j.ydbio.2023.09.011.

In Vivo Tracing of the Cytokeratin 14 Lineages using Self-Cleaving guide RNAs and CRISPR/Cas9

Amita Tiyaboonchai[*,1,2], **Leslie Wakefield**[1,2], **Anne Vonada**[1,3], **Catherine L. May**[4,5], **Craig Dorrell**[1,2], **David Enicks**[1,2], **Anusha Sairavi**[3], **Klaus H. Kaestner**[4,5], **Markus Grompe**[1,2,3]

[1]Oregon Stem Cell Center, Oregon Health & Science University; Portland, OR 97239, USA

[2]Department of Pediatrics, Oregon Health & Science University; Portland, OR 97239, USA

[3]Department of Molecular and Medical Genetics, Oregon Health & Science University; Portland, OR 97239, USA

[4]Institute of Diabetes, Obesity, and Metabolism, Perelman School of Medicine, The University of Pennsylvania; Philadelphia, PA 19104, USA

[5]Department of Genetics, Perelman School of Medicine, The University of Pennsylvania; Philadelphia, PA 19104, USA

Abstract

The current gold-standard for genetic lineage tracing in transgenic mice is based on cell-type specific expression of Cre recombinase. As an alternative, we developed a cell-type specific CRISPR/spCas9 system for lineage tracing. This method relies on RNA polymerase II promoter driven self-cleaving guide RNAs (scgRNA) to achieve tissue-specificity. To demonstrate proof-of-principle for this approach a transgenic mouse was generated harbouring a knock-in of a scgRNA into the Cytokeratin 14 (Krt14) locus. Krt14 expression marks the stem cells of squamous epithelium in the skin and oral mucosa. The scgRNA targets a Stop cassette preceding a fluorescent reporter in the Ai9-tdtomato mouse. Ai9-tdtomato reporter mice harbouring this allele along with a spCas9 transgene demonstrated precise marking of the Krt14 lineage. We conclude that RNA polymerase II promoter driven scgRNAs enable the use of CRISPR-/spCas9 for genetic lineage tracing.

INTRODUCTION

Lineage tracing is important for the understanding of stem cell fate, tissue development and disease in vivo. In lineage tracing, a single cell or group of cells are marked in a heritable way such that the fate of their progeny can then be observed. Various methods have been used for lineage tracing, including direct observation, vital dye or radioactive tracer labelling and transduction of marker genes by viral vectors. However, the most commonly

[*]Corresponding author. tiyaboon@ohsu.edu.

CONFLICT OF INTEREST

Dr. Grompe has significant financial interests in Yecuris Corp., a company that may have commercial interest in the results of this research.

used tool for lineage tracing today employs marker gene activation by cell-type specific Cre recombination of loxP sites (Das and Yaniv, 2020; Kretzschmar and Watt, 2012; Liu et al., 2020).

In the Cre-loxP system, Cre recombinase recognizes 34 base pair loxP sites and catalyses the recombination of the DNA between two loxP sites (Kim et al., 2018). Dependent on the directionality of the loxP sites, a deletion or an inversion of the flanked region occurs. The development of transgenic mice with cell type specific promoters driving Cre recombinase expression allowed temporal and spatial control of recombination in living mice (Kim et al., 2018). Additionally, tamoxifen-inducible forms of Cre have been developed that allow drug-induced conditional recombination. Cre-ER is a version of Cre recombinase containing a mutated ligand binding domain of the estrogen receptor fused to the Cre recombinase, resulting in Cre becoming specifically activated by tamoxifen (Feil et al., 2009; Hayashi and McMahon, 2002).

Genetic lineage tracing is most commonly performed using a Cre-loxP system in which the Cre recombinase is driven by a cell type-specific RNA polymerase II (RNAP II) promoter (He et al., 2017). While the Cre-loxP system is a powerful tool for lineage tracing and in vivo genetic modification, there are limitations to its use. Cre reporters can be leaky (Álvarez-Aznar et al., 2020; Liu et al., 2010; Schmidt et al., 2000; Vooijs et al., 2001). Cre can also be inefficient, especially in weakly expressed genes. Additionally, the Cre recombinase protein is large and high expression levels result in toxicity in neonatal mice (Forni et al., 2006), cells that rapidly proliferate (Kurachi et al., 2019) and cardiomyocytes (Pugach et al., 2015). Expression of the Cre recombinase protein in developing spermatids can also result in male sterility due to chromosomal rearrangements in the spermatids (Schmidt et al., 2000). While the Cre-ER fusion protein provides temporal control over the lineage tracing markings that are desired, tamoxifen has significant toxicity (Donocoff et al., 2020; Vasioukhin et al., 1999) and is especially poorly tolerated in developing embryos (Green et al., 2005; Sun et al., 2021).

Here we present a cell type specific CRISPR/spCas9 system using a scgRNA strategy as an alternative method to Cre recombinase for in vivo lineage tracing in a mouse model. The Ai9-tdTomato mouse contains a tdTomato marker gene that can be activated by CRISPR/spCas9 mediated removal of a translational Stop cassette downstream of a promoter (Staahl et al., 2017). However, guide RNAs (gRNAs) are typically driven by RNA polymerase III promoters which are active ubiquitously and hence do not allow cell type specificity. To be useful for lineage tracing in a similar way as Cre recombinase, a polymerase II (RNAP II) promoter driven construct is necessary. We and others have recently demonstrated that self-cleaving guide RNAs (scgRNA) can allow a functional gRNA to be expressed from an endogenous RNAP II promoter in vivo drosophila and in mice (Garcia-Marques et al., 2019; Tiyaboonchai et al., 2022). A scgRNA is a gRNA flanked by two ribozymes, the hammerhead and hepatitis delta ribozymes (Gao and Zhao, 2014; Lee et al., 2016). Upon transcription from the DNA, the ribozymes fold into their secondary structure resulting in self catalytic activity releasing the functional gRNA. This previously developed strategy of expressing scgRNAs from a polymerase II promoter is an ideal candidate for use in lineage tracing. A potential advantage of an RNAP II driven scgRNA for lineage tracing is their

much smaller size when compared to Cre recombinase. In addition, we have shown that functional scgRNAs can be produced from weak RNAP II promoters and as a result may have a superior dynamic range compared to the Cre-loxP system (Tiyaboonchai et al., 2022).

To test this alternative method for genetic marking, Cytokeratin 14 (Krt14) was chosen because it is a well-known epithelial stem cell marker found throughout the body including in the salivary glands (Kwak et al., 2016), basal cells of the skin (Blanpain and Fuchs, 2006; Coulombe et al., 1989), and the tongue (Okubo et al., 2009). Additionally, Krt14 has a well characterized tissue distribution pattern (Hafner et al., 2004; Vasioukhin et al., 1999; Wang et al., 1997). Here, a transgenic mouse was generated harbouring a knock-in of a scgRNA into the mouse Krt14 locus. This scgRNA targets the Stop sequence that prevents reporter gene expression in the Ai9-tdTomato mouse. In transgenic mice bearing this knock-in as well as a spCas9 transgene and the Ai9 reporter, tdTomato was found to be expressed in known Krt14 expressing tissues (Hafner et al., 2004; Wang et al., 1997), indicating tightly specific scgRNA expression from the Krt14 promoter. When a mouse with a doxycycline-controllable spCas9 was used, doxycycline administration allowed temporal control over the marking. Together these data demonstrate the ways in which scgRNAs can be utilized to allow the specific expression of a CRISPR/spCas9 gRNA from a RNAP II promoter in vivo.

MATERIAL AND METHODS

Animal Husbandry

All animal experiments were approved by the Institutional Animal Care and Use Committee at the Oregon Health & Science University. All animal work was performed according the guidelines set for animal care at the Oregon Health & Science University. B6-Rosa-Ai9tdTomato (stock #007909), B6J-Rosa-Cag-spCas9 (stock #026179), B6-Cg-Col1a1-TRE-Cas9 (stock #029476) and B6-Rosa-rtTA (stock # 006965) were obtained from Jackson Laboratories. The mice were fed standard mouse chow (LabDiet Piclab Rodent Diet 5LOD) or standard mouse chow containing 625 mg/kg Doxycycline (Envigo TD.01306) and tap water. Doxycycline inducible experimental mice were placed on the doxycycline mouse chow for 2-4 weeks at a time. All mice were house in groups of 2 to 4 animals in Thoren caging under standard 12 hour on and off light cycles in the Department of Comparative Medicine at the Oregon Health & Science University. The animal rooms were maintained at 70F and 30%-70% humidity.

To generate experimental animals homozygous B6-Rosa-Ai9tdTomoato mice were bred to homozygous B6J-Rosa-Cag-spCas9 mice. The F1 offspring were then bred to the Krt14-scgRNA mouse. Mice were genotyped for the Krt14-scgRNA transgene using the primers: forward, 5’-TCTCTCCCCTTCTCCAGTGA-3’ and reverse, 5’-GGGACAATACAGGGGCTCTT-3’. The presence of the transgene results in an amplicon size of 403 base pairs. To generate doxycycline-mediated inducible K14-scgRNA experimental animals, we generated and maintained two holdings that would be bred together to create the target genotype. The first line consisted of mice homozygous for both the B6-Rosa-Ai9tdTomato and B6-Cg-Col1a1-TRE-Cas9 transgenes. The second line

consist of mice homozygous for B6-Rosa-rtTA and Krt14-scgRNA transgenes. Mice from these two lines were then bred together to generate the desired genotype.

Generation of the Krt14-scgRNA mouse

The Krt14-scgRNA mouse was generated by pronuclear injection of a single stranded repair DNA template introducing the Krt14-scgRNA transgene and a Krt14 gRNA in a CRISPR/spCas9 system. The Krt14 gRNA (5'-TGCGTGGACGCAGATACTGC-3') was designed using CRISPOR (http://crispor.tefor.net) (Haeussler et al., 2016) to target the terminal exon (exon 8) of murine Krt14 in the 3'UTR and synthesized as a crRNA (Integrated DNA Technologies). The single-stranded repair template contains arms of homology to the terminal exon of murine Krt14 (chr11:100203162-100203353 in the mouse genome assembly GRCm38/mm10). These arms of homology flank a scgRNA consisting of a hammerhead ribozyme, followed by a Stop-gRNA (5'-AAACCTCTACAAATGTGGTA-3') that targets the Stop cassette in the Ai9-tdTomato mouse(Staahl et al., 2017), the spCas9 tracr RNA scaffold and then a hepatitis delta virus ribozyme. The sequence of the transgene can be found in Table 1

To generate ribonuclear proteins (RNPs), 5 μg of crRNA and 10 μg of Alt-R® CRISPR-Cas9 tracrRNA (Integrated DNA technologies, 1072534) were annealed in a thermocycler. 2 μg of the annealed product and 500 ng of Alt-R® S.p. HiFi Cas9 Nuclease V3 (Integrated DNA Technologies, 1081058) were combined with injection buffer to a total volume of 90 μl and incubated at room temperature for 15 minutes. 1 μg at a concentration of 100 ng/μL of the single stranded repair template was then added to the mixture. The entire 100 μL mixture was sterile filtered through a 0.22 μm pore size membrane filter (Millipore Sigma, SLGV013SL). The mixture was stored at −20°C until the date of injection. The gRNA and single stranded repair template were delivered via pronuclear injection into oocytes and implanted into pseudopregnant foster mothers.

The presence of the transgene was confirmed in pups by PCR using the following primer pair which flank the region where Stop-scgRNA transgene should be inserted: forward, 5'-TCTCTCCCCTTCTCCAGTGA-3' and reverse, 5'-GGGACAATACAGGGGCTCTT-3'. The presence of the transgene results in an amplicon size of 403 base pairs. Standard PCR was performed using MyTaq Red master mix (Meridian Bioscience, BIO-25044) with an annealing temperature of 56°C. This same PCR was used for genotyping of animals subsequently generated from the founder Krt14-scgRNA mouse. Additionally, the integration of the Krt14-scgRNA transgene into the Krt14 locus was confirmed by PCR amplification of the transgenes with primers located outside of the arms of homology in the repair template followed by Sanger sequencing of the amplicon (Supplemental Figure 1). The primers used for this amplicon are: forward, 5'-GGAGTGAGGTGGTAAACGGG-3' and reverse, 5'-TCTGAGGTGGTGAAACCGAC-3'.

Southern blot

Genomic DNA was isolated from spleens of mice using the MasterPure Complete DNA & RNA extraction kit (Epicenter) following the manufacturer's protocol. 10 μg of DNA was digested overnight with the restriction enzymes *BglII* and *NdeI* (New England Biolabs).

The digested DNA was then electrophoresed in a 1.5% agarose gel in 1xTAE buffer at 100 volts for 240 minutes. The gel was transferred to a glass baking dish and soaked in alkaline transfer buffer (0.4 M NaOH, 1 M NaCl) for 20 minutes, 2 times. An overnight (18 hours) gel transfer was set with a charged nylon Zeta probe membrane (BioRad) and alkaline transfer buffer. The following day, the transfer setup was diassembled and the membrane was placed in neutralization buffer (0.5 M Tris-Cl, 1 M NaCl) for 15 minutes. The membrane was briefly washed with 2x sodium saline citrate and cross-linked using a UV cross linker (Spectroline). The membrane was blocked with Hybridization Buffer (final: 6X SSC, 0.01% Tris-HCl pH 8.0, 1% SDS, 5X Denhardt's) for 30 minutes at 65°C. Gel-purified DNA containing the probe region of interest (the Stop scgRNA) was labelled with ^{32}P using Prime-It II Random Primer Labelling kit (Agilent Technologies) and membrane was hybridized with the probe overnight in a rotor oven at 65C. The membrane was washed with wash buffer (0.1% SDS and 0.1x SSC) in the rotor oven for a total of 8 times (twice overnight). The membrane was visualized using Amersham Typhoon and analysed with ImageQuant software.

Histology and imaging

Upon harvest, tissue was fixed in 4% paraformaldehyde (Sigma-Aldrich, P64148) in 1x Dulbecco's phosphate buffered saline (DPBS) for 4 hours at room temperature or overnight at 4˚C. Tissues were prepared for cryopreservation by passing through a 10%, 20% and 30% sucrose (Thermo Fisher Scientific, BP220-212) in DPBS (w/v). Tissues were then embedded in Tissue-Tek optimal cutting temperature compound (Thermo Fisher Scientific, 4583) and cut into 7-8 μm sections using a cryostat onto Colorfrost Plus microscope slides (Thermo Fisher Scientific, 1255019). For immunofluorescence staining, tissue sections were permeabilized with 0.25% Triton X-100 in PBS at room temperature for 12 minutes then washed in DPBS 3x5 minutes. Tissue sections were blocked in 10% normal donkey serum for 30 minutes at room temperature. Slides were incubated with the primary antibody, rabbit-anti-Cytokeratin14 (Invitrogen, cat #MA5-16370) at room temperature for 1 hour followed by 3x5 minute wash in DPBS. Slides were incubated in the secondary antibody, donkey-anti-rabbit-A647 (Invitrogen, cat#A32795) for 1 hour at room temperature or overnight at 4°C followed by 3x5 minute was in DPBS. Prior to imaging, all tissue sections were counterstained with Hoechst 33342 (Invitrogen, H3570) at 1:10,000 dilution in 1xDPBS for 3 minutes then washed 3x5 minutes in 1xPBS. Coverslips were mounted with Southern Biotech Fluoromount-G slide mounting medium (Thermo Fisher Scientific, OB100-01). Imaging was performed on a Zeiss LSM700 confocal microscope.

Tissues from a total of 7 male and 8 female mice that had constitutively active spCas9 (Krt14-scgRNA; spCas9; Ai9tdTomato) were observed for the presence of tdTomato. These mice consisted of both littermates from the same breeding pair as well as non-littermates of the same genotype from independent breeding pairs. Tissues from 3 mice with the genotype Krt14-scgRNA;spCas9 and 4 mice with the genotype spCas9;Ai9tdTomato were analyzed. Tissues from 4 Krt14-scgRNA; spCas9; Ai9tdTomato mice were stained for Krt14. For the doxycycline inducible spCas9 containing mice, 6 mice that were maintained on doxycycline diet for 2 weeks and 4 mice that were maintained on the diet for 4 weeks were analyzed.

Tissues from 6 mice containing a doxycycline inducible spCas9, were maintained on a doxycycline diet for 2 weeks and 4 mice containing a doxycycline inducible spCas9 were maintained on doxycycline diet for 4 weeks were analysed for tdTomato.

Isolation of tongue epithelial cells

Tongues excised from freshly sacrificed mice were trimmed with scissors to recover the outer epithelium and exclude most muscle tissue. The resulting chunks of epithelial tissue were transferred to a solution of 0.2% dispase II (Thermo Fisher), 1.5 mg/ml Type 4 collagenase (Worthington), and 0.1% DNase I (Sigma), minced with a razor blade, and stirred at 37°C for 1.5 hours with occasional pipetting to assist tissue dispersal. Cells were then collected through a 40 μm strainer and labelled with APC conjugated anti-CD45 (BD Pharmingen) and propidium iodide (Sigma) for dead cell exclusion. Using a BD Influx cell sorter, live (PI-) non-hematopoietic (CD45-) cells were sorted to recover tdTomato+ and tdTomato− fractions and flash frozen for molecular analysis. Genomic DNA was isolated from the sorted cells as well as a control spleen from a Krt14-scgRNA; Rosa-Ai9tdTomato mouse using the MasterPure Complete DNA & RNA extraction kit (Epicenter). Standard PCR was performed using MyTaq Red master mix (Meridian Bioscience, BIO-25044) with an annealing temperature of 60°C for 40 cycles. Primers for the amplification of the Stop cassette region used were: forward 5’-GCTCCTGGGCAACGTGCTGGTTATTG-3’ and reverse 5’-TTGATGACCTCCTCGCCCTTGCTCAC-3’ resulting in an amplicon that is 1130 base pairs in size when the Stop cassette has not been manipulated. Resulting PCR amplicons were run on a 1.2% agarose gel. PCR products were cloned into the pCR4-TOPO vector using the TOPO TA Cloning for sequencing kit (Invitrogen) and Sanger sequenced. Sequences were analysed using Snapgene software (from Dotmatics; available at snapgene.com).

RESULTS

Generation of the Krt-14 scgRNA mouse

The Keratin14-scgRNA (Krt14-scgRNA) lineage tracing mouse was created by using a scgRNA targeting the Stop cassette (Stop-scgRNA) (Tiyaboonchai et al., 2022) of the Ai9-tdTomato mouse. The Ai9-tdTomato mouse harbours a transgene in the *Rosa26* locus which consists of a CAG promoter followed by a Stop cassette-controlled tdTomato fluorescent reporter (Madisen et al., 2010). Untreated, this mouse does not express the fluorescent tdTomato reporter in any cells due to the presence of the Stop cassette. The Stop cassette consists of three repetitive sequences that can be inactivated by targeting with a gRNA in a CRISPR/Cas9 system (Staahl et al., 2017). When the Stop cassette is lost or inactivated, the CAG promoter can drive the expression of tdTomato, a bright fluorescent reporter. To create a Krt14-scgRNA lineage tracing mouse, the Stop-scgRNA was knocked into the 3-’UTR of the Krt14 locus downstream of the terminal coding exon by homologous recombination in zygotes (Figure 1). Sanger sequencing confirmed that the knock in transgene did not cause any mutations in the area surrounding the targeted DNA (Figure S1A) and a Southern blot demonstrated that there was a single integration of the construct into the genome (Figure S1B).

The resulting founder Krt14-scgRNA mouse was bred to animals that constitutively express spCas9 and Ai9-tdTomato from the *Rosa26* locus (Madisen et al., 2010; Platt et al., 2014). Tight expression of the reporter gene activating the scgRNA in only Krt14-expressing cells will result in tdTomato reporter activity in only the Krt14 lineages. Adult mice expressing all three transgenes were terminally harvested and various tissues were analysed for the expression of the tdTomato reporter. TdTomato-positive (tdTomato+) cells were observed in the skin, tongue, salivary glands, thymus, oesophagus and forestomach (Figure 2A and Figure S2). No tdTomato+ cells were observed in the ovaries or the testis (Figure 2A). While there have been reports of Krt14 expression in the ovaries (Hafner et al., 2004), immunohistochemistry and RNA sequencing references in the Human Protein Atlas show no detection of Krt14 expression in the ovaries (Digre and Lindskog, 2021; "Human Protein Atlas"; Uhlen et al., 2015). In the Krt14-scgRNA mice, tdTomato+ cells were not observed in the ovaries or the testis (Figure 2A). In the submandibular salivary glands, clusters of tdTomato+ cells were observed (Figure 2A). Krt14 is expressed predominantly in the progenitor stem cells of the salivary glands and cells decline in number by 15-fold between 2 to 8 weeks of age (Kwak et al., 2016) becoming a rare population in adult mice. In the Krt14-scgRNA mouse, when the tdTomato reporter expression is induced, it is permanent in the cell and its daughter cells most likely leading to the presence of the population of tdTomato+ cells in the salivary glands of these adult mice. Additional tissues including the kidney, liver, spleen, pancreas and bladder (Figure S2) were also imaged and observed to be completely tdTomato negative (tdTomato−) as expected, demonstrating that the lineage tracing system is not leaky in other tissues. In the tdTomato+ tissues, the tdTomato+ expression was noticed to not be uniform across the tissues. This was found to be consistent across littermates and mice of the same genotype from independent parents (Figure S2). Additionally, mice that were bred to either lack spCas9 (only contains Krt14-scgRNA and Ai9-tdTomato) or lack Krt14-scgRNA (only contains Ai9-tdTomato and spCas9) did not have any tdTomato+ cells present in the tongue, skin, salivary glands, forestomach or esophagus (Figure 2B, 2C and Figure S4) further confirming the precise expression of the reporter.

Immunofluorescence (IF) staining against Krt14 was done in a subset of tissues from the Krt14-scgRNA mouse. The majority of the cells that were tdTomato+ within the tongue, skin, and forestomach were colocalized with the Krt14 staining. The ovaries did not stain positive for Krt14 (Figure 2D and Figure S3). Tissues from both littermates and mice of the same genotype from independent parents were also stained for Krt14 and confirmed that these phenotypes were consistent across the animals (Figure S3). These data show that the Krt14-scgRNA knock-in mouse can mark Krt14 expression previously known from other lineage tracing systems (Hafner et al., 2004; Vasioukhin et al., 1999; Wang et al., 1997).

In all of the tissues positive for Krt14 by IF staining, cells that did not express the reporter (tdTomato− cells) while positive for Krt14+ by IF were also observed. One possible explanation for the presence of these cells is that not all CRISPR-induced double-stand breaks result in activation of the reporter. Once the gRNA cleaves the DNA resulting in a double stand break, it is most commonly repaired by non-homologous end joining (NHEJ) (Xue and Greene, 2021). This imprecise repair may result in mutations that are insufficient to allow tdTomato expression, but sufficient to disrupt the gRNA recognition

sequence and prevent recutting. To investigate this possibility, epithelial cells of the tongue, a known Krt14+ population, were subjected to fluorescence activated cell sorting (FACS) for tdTomato− and tdTomato+ cells (Figure S5). DNA was isolated from both of the populations of sorted cells as well as from control cells with an unmanipulated Stop-cassette region (derived from a Krt14-scgRNA; Ai9-tdTomato mouse lacking spCas9). The region of the Stop cassette was PCR amplified and visualized on an agarose gel (Figure 3A). The control cells had one amplicon consistent in size with the full length unmanipulated Stop cassette. The tdTomato− cells had multiple amplicons of varying sizes, and the tdTomato+ cells contain a single, smaller PCR product. These PCR products were TOPO cloned and Sanger sequenced (Figure 3B). As expected, PCR amplicons from the control cells contained the full-length Stop cassette without any deletions (5/5 analyzed). PCR amplicons from the tdTomato+ cells displayed a loss of at least two of repetitive sequences in the Stop cassette with a subset containing an additional indel close to the PAM site of the third Stop cassette (5/5 analyzed). In the tdTomato− cells, however, there were three different types of amplicons. Some of the amplicons contained a loss of only a single repetitive sequence, some contained small deletions disrupting the gRNA target sequence and some contained a combination of these large and small deletions (7/10 analyzed). Lastly, there was also a subset of the amplicons which had the full-length, unmanipulated Stop cassette (3/10 analyzed). These results demonstrate that the population of tdTomato-, Krt14+ cells is predominantly due to the disruption of the gRNA targeting sequence but not a loss of the Stop cassette during NHEJ or insufficient loss of copies of the repetitive sequence.

Doxycycline inducible Krt14-scgRNA

To further explore the use of the scgRNA as a tool for lineage tracing, it is valuable to have temporal control over the gene marking process. When using a Cre-loxP system, conditional activation of Cre can be achieved by use of the Cre recombinase fused to the human estrogen receptor (Cre-ER) and the addition of tamoxifen. Various modifications have been made to Cre-ER to prevent leakiness and increase sensitivity to tamoxifen. However, this resulted in the need of high amounts of tamoxifen for activation which is sometimes toxic (Kretzschmar and Watt, 2012; Vasioukhin et al., 1999). Tamoxifen toxicity may be especially problematic during early development and embryogenesis (Sun et al., 2021). In contrast, temporal control of spCas9 could be achieved with doxycycline-inducible spCas9 (Bowling et al., 2020).

The Krt14-scgRNA mouse was bred to co-express an inducible spCas9 controlled by the tetracycline response element (TRE-spCas9), the reverse tetracycline response activator (rTTA) and Ai9-tdTomato. When doxycycline is administered, the TRE promoter can then drive the expression of the downstream transgene, spCas9. To induce spCas9, adult mice were placed on doxycycline diet for 2 or 4 weeks. Tissues from the mice were collected and imaged for tdTomato. In agreement with the mice that constitutively express spCas9, these mice also had tdTomato+ cells in the tongue, skin, salivary glands and oesophagus although at varying lower frequencies (Figure 4 and Figure S6). The thymus and forestomach were observed to be tdTomato− (Figure 4 and Figure S6). In the skin and tongue of mice that had been maintained on a doxycycline diet for 4 weeks, the frequency of tdTomato+ cells was higher than those that had been maintained on a doxycycline diet for 2 weeks. In the

salivary glands and esophagus, tdTomato+ cells were rare and observed at a similar rate at both timepoints. Additional tissues including the liver, kidney, pancreas, spleen and bladder were imaged and found to be tdTomato− negative (Figure S7). Age matched littermates co-expressing all four transgenes that were kept on regular chow without doxycycline did not express tdTomato in any of the tissues that were examined (Figure S8). Additionally, littermates lacking the rTTA transgene that were maintained on doxycycline diet for 4 weeks did not express tdTomato in any of the tissues imaged (Figure S8). These results demonstrate tight regulation of the Stop-sgRNA following induction of spCas9 expression by doxycycline.

DISCUSSION

Together these data establish that knocking scgRNAs into a RNAP II-driven cell-type specific gene is a viable approach for creating transgenic lineage tracing lines. We have shown that the Krt14-scgRNA in a mouse constitutively expressing spCas9 and Ai9-tdTomato results in tight expression of the tdTomato reporter in tissues with known Krt14 expression. Krt14 is marker of epithelial stem cells, and the Krt14-scgRNA mouse generated here can be used for lineage tracing of these progenitor cells in various tissue types.

Throughout the tissues, tdTomato expression was not as uniform as endogenous Krt14 expression. By cells sorting tdTomato+ and tdTomato− tongue epithelial cells we were able to demonstrate that the majority of tdTomato− cells that are Krt14+ result from a combination of a deletion of only a single repetitive sequence of the Stop cassette or indel mutations that do not cause a sufficient deletion of the Stop cassette. This result indicates that the lack of complete marking is due to the design of the reporter locus, not the lack of CRISPR activity. There was also a small subset of tdTomato− cells that retained an unmanipulated Stop cassette. This likely results from contaminating cells (our FACS procedure can also capture some non-epithelial cells) or a subset of cells having a closed chromatin structure making it difficult for access by the CRISPR-spCas9 complex. In future, development of a reporter gene that is dependent on only a single gRNA cleavage or cherry-stringing multiple scgRNAs that have different targets in the Stop-cassette would promise to yield substantially higher marking efficiency. Cells which were tdTomato+ and Krt14− were also observed in some of the tissues. Advantageously for lineage tracing, tdTomato expression is permanent once the stop cassette has been removed. Hence, all cells that have expressed Krt14 and subsequently their progeny will continue to be tdTomato+ even if there is no longer Krt14 expression in that cell.

While we demonstrate here that the scgRNA system can be utilized for activation of a reporter, scgRNA transgenes could also be targeted to any somatic locus to create a cell-type specific gene knockout by DNA double-strand breaks. In contrast to the Cre-loxP system, the construction of a floxed allele at the target locus would not be necessary. Additionally, scgRNAs could be used to achieve cell-type specific gene activation or silencing in transgenic mice expressing nuclease-dead versions of spCas9 fused to transcriptional activators or repressor. Thus, the generation of a single scgRNA transgenic mouse could be utilized not only for lineage tracing but also for the functional study of gene regulation.

The use of RNAP II-driven scgRNAs has powerful potential in transgenesis. Many Cre reporters are inefficient when targeting weakly expressed genes. Our prior work suggests that scgRNA reporters have a wide dynamic range and hence they may work better for such loci (Tiyaboonchai et al., 2022). Incorporation of the scgRNA into any progenitor cell of interest within an inducible spCas9 system should be feasible and would allow the delivery of a pulse of doxycycline for lineage tracing of the resulting progeny. Due to the low toxicity of doxycycline, this can be done during embryogenesis. Although different genes have different expression levels, the activation of tdTomato only relies upon a minimum of 2 double strand breaks made by the CRISPR/Cas9 system. Furthermore, scgRNA cassettes are less than 300 base pairs in size, much smaller than Cre recombinase, facilitating the construction of transgenic lines. Finally, conditional activation of Cre-ER to achieve temporal control of marking requires tamoxifen, which is toxic (Vasioukhin et al., 1999) and problematic to use in embryogenesis (Sun et al., 2021). We have demonstrated that the activation of a fluorescent reporter, tdTomato, can be temporally controlled by expression of a scgRNA from a polymerase II promoter multiplexed with a doxycycline-inducible spCas9 system. Following maintenance on a doxycycline diet, the skin, tongue and salivary glands were well labelled. There was consistently sparse labelling of the esophagus and no labelling observed in the thymus and forestomach in these tissues demonstrated to have lower expression of Krt14 as compared to the skin and tongue (Wang et al., 1997). A higher dose of doxycycline may be required for sufficient and consistent expression of the spCas9 to improve reporter expression in these tissues although sparse labelling of cells within a tissue can be advantageous for lineage tracing the progeny of a specific cell. While the system described here can overcome some of the limitations of Cre-ER, the generation of the Krt14-scgRNA doxycycline-inducible mice may be fairly time consuming due to the number of transgenes that must be incorporated into a single mouse.

We have demonstrated the feasibility of the scgRNA for lineage tracing in the Krt14 locus in both a constitutive and inducible manner. This strategy can be expanded to lineage trace any cells with known specific gene expression in a tissue by knocking in the Stop-scgRNA into any polymerase II driven endogenous loci. While genes with weak expression may pose a challenge, an improvement to transgene reporter induction may be possible by cherry-stringing 2-3 different scgRNAs that all target the Stop-cassette.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We thank the transgenic and chimeric mouse core facility at the University of Pennsylvania Perelman School of Medicine for the generation of the founder Keratin14 scgRNA mouse. We would also like to thank Hiroyuki Nakai’s lab at OHSU for resources and assistance with the radioactive portions of the Southern blot.

FUNDING

This work was supported by the Ruth L. Kirschstein National Research Service Awards Postdoctoral Fellow from the National Institute of Diabetes and Digestive and Kidney Diseases (grant number: 1 F2 DK117516-01). This research was sponsored in part by a research agreement with LogicBio Therapeutics.

REFERENCES

Álvarez-Aznar A, Martínez-Corral I, Daubel N, Betsholtz C, Mäkinen T, Gaengel K, 2020. Tamoxifen-independent recombination of reporter genes limits lineage tracing and mosaic analysis using CreERT2 lines. Transgenic Res. 29, 53–68. 10.1007/s11248-019-00177-8 [PubMed: 31641921]

Blanpain C, Fuchs E, 2006. Epidermal stem cells of the skin. Annu. Rev. Cell Dev. Biol 10.1146/annurev.cellbio.22.010305.104357

Coulombe PA, Kopan R, Fuchs E, 1989. Expression of keratin K14 in the epidermis and hair-follicle: Insights into complex programs of differentiation. J. Cell Biol 109, 2295–2312. 10.1083/jcb.109.5.2295 [PubMed: 2478566]

Das RN, Yaniv K, 2020. Discovering new progenitor cell populations through lineage tracing and in vivo imaging. Cold Spring Harb. Perspect. Biol 12, 1–25. 10.1101/cshperspect.a035618

Digre A, Lindskog C, 2021. The Human Protein Atlas—Spatial localization of the human proteome in health and disease. Protein Sci. 30. 10.1002/pro.3987

Donocoff RS, Teteloshvili N, Chung H, Shoulson R, Creusot RJ, 2020. Optimization of tamoxifen-induced Cre activity and its effect on immune cell populations. Sci. Rep 10, 1–12. 10.1038/s41598-020-72179-0 [PubMed: 31913322]

Feil S, Valtcheva N, Feil R, 2009. Inducible cre mice, Methods in Molecular Biology. 10.1007/978-1-59745-471-1_18

Forni PE, Scuoppo C, Imayoshi I, Taulli R, Dastrù W, Sala V, Betz UAK, Muzzi P, Martinuzzi D, Vercelli AE, Kageyama R, Ponzetto C, 2006. High levels of Cre expression in neuronal progenitors cause defects in brain development leading to microencephaly and hydrocephaly. J. Neurosci 26, 9593–9602. 10.1523/JNEUROSCI.2815-06.2006 [PubMed: 16971543]

Gao Y, Zhao Y, 2014. Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. J. Integr. Plant Biol 56, 343–349. 10.1111/jipb.12152 [PubMed: 24373158]

Garcia-Marques J, Yang CP, Espinosa-Medina I, Mok K, Koyama M, Lee T, 2019. Unlimited Genetic Switches for Cell-Type-Specific Manipulation. Neuron 104, 227–238.e7. 10.1016/j.neuron.2019.07.005 [PubMed: 31395429]

Green AR, Styles JA, Parrott EL, Gray D, Edwards RE, Smith AG, Gant TW, Greaves P, Al-Azzawi F, White INH, 2005. Neonatal tamoxifen treatment of mice leads to adenomyosis but not uterine cancer. Exp. Toxicol. Pathol 56, 255–263. 10.1016/j.etp.2004.10.001 [PubMed: 15816354]

Haeussler M, Schönig K, Eckert H, Eschstruth A, Mianné J, Renaud JB, Schneider-Maunoury S, Shkumatava A, Teboul L, Kent J, Joly JS, Concordet JP, 2016. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome Biol. 17, 1–12. 10.1186/s13059-016-1012-2 [PubMed: 26753840]

Hafner M, Wenk J, Nenci A, Pasparakis M, Scharffetter-Kochanek K, Smyth N, Peters T, Kess D, Holtkötter O, Shephard P, Kudlow JE, Smola H, Haase I, Schippers A, Krieg T, Müller W, 2004. Keratin 14 Cre Transgenic Mice Authenticate Keratin 14 as an Oocyte-Expressed Protein. Genesis 38, 176–181. 10.1002/gene.20016 [PubMed: 15083518]

Hayashi S, McMahon AP, 2002. Efficient recombination in diverse tissues by a tamoxifen-inducible form of Cre: A tool for temporally regulated gene activation/inactivation in the mouse. Dev. Biol 244, 305–318. 10.1006/dbio.2002.0597 [PubMed: 11944939]

He L, Li Yan, Li Yi, Pu W, Huang X, Tian X, Wang Y, Zhang H, Liu Q, Zhang L, Zhao H, Tang J, Ji H, Cai D, Han Zhibo, Han Zhongchao, Nie Y, Hu S, Wang QD, Sun R, Fei J, Wang F, Chen T, Yan Y, Huang H, Pu WT, Zhou B, 2017. Enhancing the precision of genetic lineage tracing using dual recombinases. Nat. Med 23, 1488–1498. 10.1038/nm.4437 [PubMed: 29131159]

Human Protein Atlas [WWW Document], n.d.. proteinatlas.org.

Kim H, Kim M, Im S-K, Fang S, 2018. Mouse Cre-LoxP system: general principles to determine tissue-specific roles of target genes. Lab. Anim. Res 34, 147. 10.5625/lar.2018.34.4.147 [PubMed: 30671100]

Kretzschmar K, Watt FM, 2012. Lineage tracing. Cell 148, 33–45. 10.1016/j.cell.2012.01.002 [PubMed: 22265400]

Kurachi M, Ngiow SF, Kurachi J, Chen Z, Wherry EJ, 2019. Hidden Caveat of Inducible Cre Recombinase. Immunity 51, 591–592. 10.1016/j.immuni.2019.09.010 [PubMed: 31618646]

Kwak M, Alston N, Ghazizadeh S, 2016. Identification of Stem Cells in the Secretory Complex of Salivary Glands. J. Dent. Res 95, 776–783. 10.1177/0022034516634664 [PubMed: 26936214]

Lee RTH, Ng ASM, Ingham PW, 2016. Ribozyme Mediated gRNA Generation for In Vitro and In Vivo CRISPR/Cas9 Mutagenesis. PLoS One 11, e0166020. 10.1371/journal.pone.0166020 [PubMed: 27832146]

Liu K, Jin H, Zhou B, 2020. Genetic lineage tracing with multiple DNA recombinases: A user's guide for conducting more precise cell fate mapping studies. J. Biol. Chem 10.1074/jbc.REV120.011631

Liu Y, Suckale J, Masjkur J, Magro MG, Steffen A, Anastassiadis K, Solimena M, 2010. Tamoxifen-independent recombination in the RIP-CreER mouse. PLoS One 5, 1–7. 10.1371/journal.pone.0013533

Madisen L, Zwingman TA, Sunkin SM, Oh SW, Zariwala HA, Gu H, Ng LL, Palmiter RD, Hawrylycz MJ, Jones AR, Lein ES, Zeng H, 2010. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat. Neurosci 13, 133–140. 10.1038/nn.2467 [PubMed: 20023653]

Okubo T, Clark C, Hogan BLM, 2009. Cell Lineage Mapping of Taste Bud Cells and Keratinocytes in the Mouse Tongue and Soft Palate. Stem Cells 27, 442–450. 10.1634/stemcells.2008-0611.Cell [PubMed: 19038788]

Platt RJ, Chen S, Zhou Y, Yim MJ, Swiech L, Kempton HR, Dahlman JE, Parnas O, Eisenhaure TM, Jovanovic M, Graham DB, Jhunjhunwala S, Heidenreich M, Xavier RJ, Langer R, Anderson DG, Hacohen N, Regev A, Feng G, Sharp PA, Zhang F, 2014. CRISPR-Cas9 knockin mice for genome editing and cancer modeling. Cell 159, 440–455. 10.1016/j.cell.2014.09.014 [PubMed: 25263330]

Pugach EK, Richmond PA, Azofeifa JG, Dowell RD, Leinwand LA, 2015. Prolonged Cre expression driven by the α-myosin heavy chain promoter can be cardiotoxic. J. Mol. Cell. Cardiol 86, 54–61. 10.1016/j.yjmcc.2015.06.019 [PubMed: 26141530]

Schmidt EE, Taylor DS, Prigge JR, Barnett S, Capecchi MR, 2000. Illegitimate Cre-dependent chromosome rearrangements in transgenic mouse spermatids. Proc. Natl. Acad. Sci. U. S. A 97, 13702–13707. 10.1073/pnas.240471297 [PubMed: 11087830]

Staahl BT, Benekareddy M, Coulon-Bainier C, Banfal AA, Floor SN, Sabo JK, Urnes C, Munares GA, Ghosh A, Doudna JA, 2017. Efficient genome editing in the mouse brain by local delivery of engineered Cas9 ribonucleoprotein complexes. Nat. Biotechnol 35, 431–434. 10.1038/nbt.3806 [PubMed: 28191903]

Sun MR, Steward AC, Sweet EA, Martin AA, Lipinski RJ, 2021. Developmental malformations resulting from high-dose maternal tamoxifen exposure in the mouse. PLoS One 16, 1–13. 10.1371/journal.pone.0256299

Tiyaboonchai A, Vonada A, Posey J, Pelz C, Wakefield L, Grompe M, 2022. Self-cleaving guide RNAs enable pharmacological selection of precise gene editing events in vivo. Nat. Commun 13, 7391. 10.1038/s41467-022-35097-5 [PubMed: 36450762]

Uhlen M, Uhlén M, Fagerberg L, Hallström BM, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Asa, Sivertsson A, Kampf C, Sjöstedt E, Sjostedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CA-K, Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist PH, Edqvist P-H, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, von Feilitzen K, von Feilitzen Kalle, Forsberg M, Persson L, Johansson F, Zwahlen M, von Heijne G, von Heijne Gunnar, Nielsen J, Ponten F, Ponten Fredrik, 2015. Proteomics. Tissue-based map of the human proteome. Science (80-.). 347.

Vasioukhin V, Degenstein L, Wise B, Fuchs E, 1999. The magical touch: Genome targeting in epidermal stem cells induced by tamoxifen application to mouse skin, Genetics.

Vooijs M, Jonkers J, Berns A, 2001. A highly efficient ligand-regulated Cre recombinase mouse line shows that LoxP recombination is position dependent. EMBO Rep. 2, 292–297. 10.1093/embo-reports/kve064 [PubMed: 11306549]

Wang X, Zinkel S, Polonsky K, Fuchs E, 1997. Transgenic studies with a keratin promoter-driven growth hormone transgene: Prospects for gene therapy. Proc. Natl. Acad. Sci. U. S. A 94, 219–226. 10.1073/pnas.94.1.219 [PubMed: 8990189]

Xue C, Greene EC, 2021. DNA Repair Pathway Choices in CRISPR-Cas9-Mediated Genome Editing. Trends Genet. 10.1016/j.tig.2021.02.008

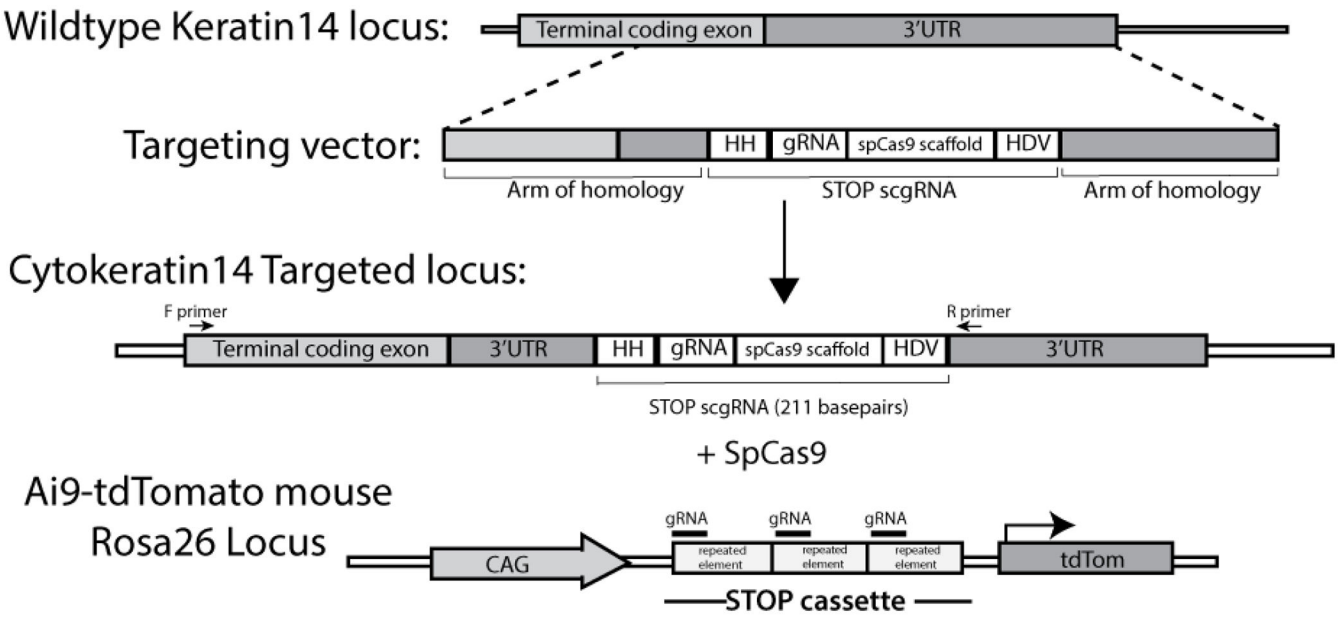

Figure 1: Targeting the Cytokeratin14 locus.

Schematic showing targeting of the Cytokeratin14 (Krt14) locus with the Stop-scgRNA into the 3'UTR following the terminal coding exon. The Ai9-tdTomato mouse contains a knock-in consisting of a CAG promoter followed by a repetitive Stop cassette downstream of tdTomato, into the Rosa26 locus. When expressed, the Stop-scgRNA causes a deletion of the Stop cassette and there is a gain-of-function of the tdTomato transgene. Arrows indicate the location of the genotyping primers. HH: Hammerhead ribozyme; gRNA: guide RNA; HDV: hepatitis delta virus ribozyme; scgRNA: self-cleaving guide RNA; CAG: chicken beta-actin promoter; tdTom: tdTomato transgene; F primer: Forward genotyping primer; R primer: Reverse genotyping primer.

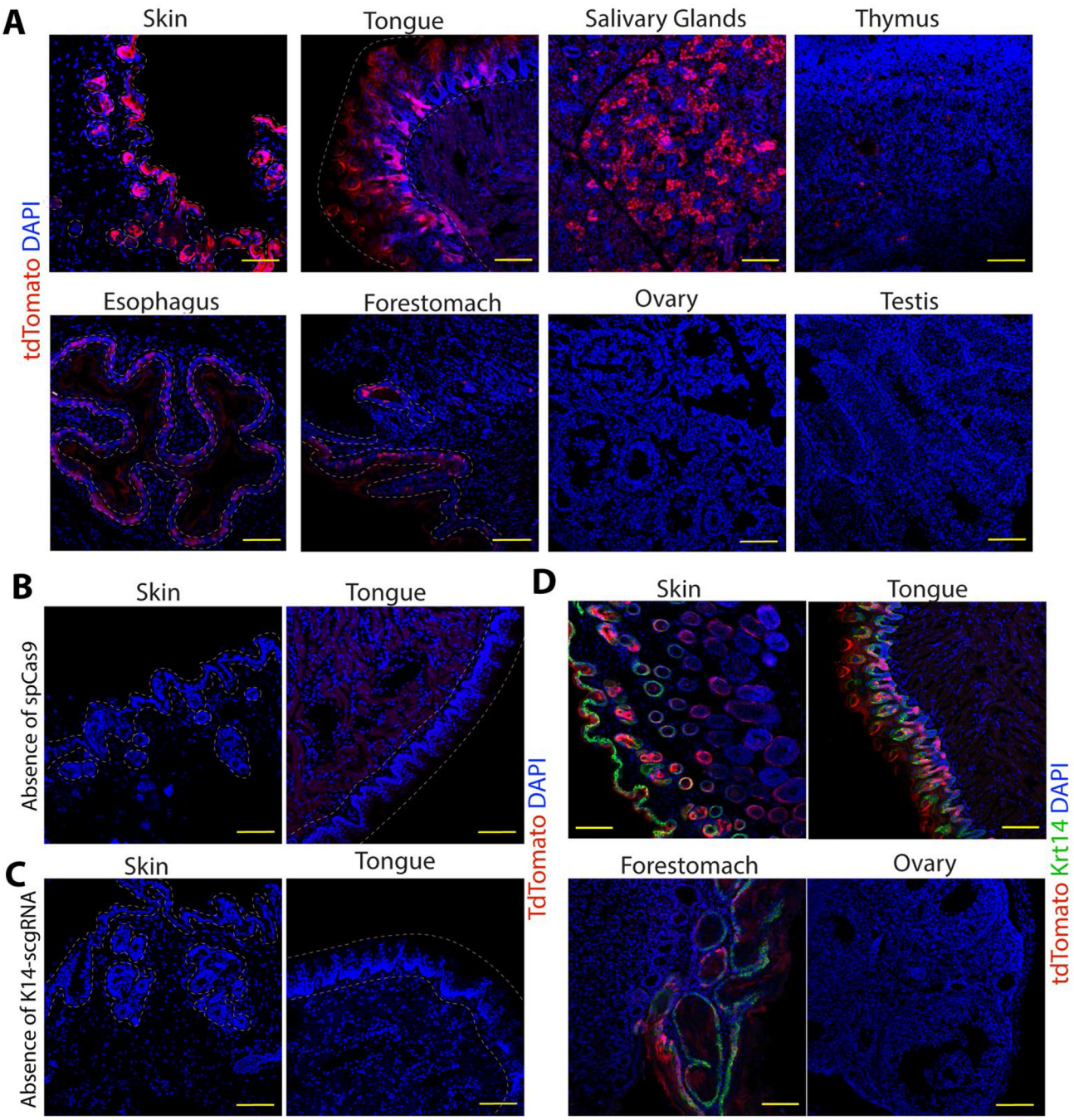

Figure 2: The Krt14-scgRNA mouse.

(A) Krt14-scgRNA mice were bred to co-express Ai9-tdTomato and constitutively express spCas9. TdTomato will only be expressed in Krt14 expressing cells due to these cells containing the knock-in scgRNA. Representative sections taken from skin, tongue, salivary glands, thymus, oesophagus, forestomach, ovary and testis of adult mice (n=15). Cells located within the dotted lines are expected to be Krt14 positive. (B) Skin and tongue from a control animal with the Krt14-scgRNA and Ai9-tdTomato transgene in the absence of the spCas9 (n=3). (C) Skin and tongue from a control animal expressing the SpCas9

and Ai9-tdTomato in the absence of Krt14-scgRNA (n=3). (D) A subset of the tissues was co-stained for Krt14 (n=4). All tissues were counterstained with Hoeschst 33342. Scale bars represent 100 μm.

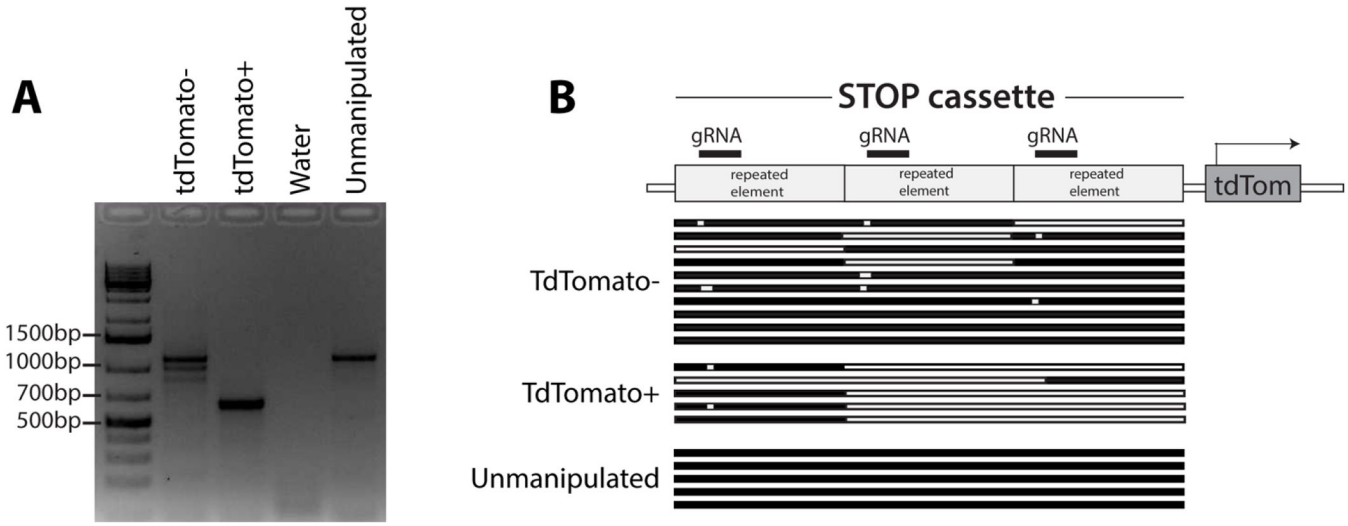

Figure 3:

(A) Agarose gel of PCR products amplified from the Stop cassette region of tdTomato− tongue basal epithelial cells (lane 2), tdTomato+ tongue basal epithelial cells (lane 3) and a control tongue from an unmanipulated Krt14-scgRNA; Rosa-Ai9tdTomato mouse (lane 5). A water control is in lane 4 and a DNA ladder is in lane 1. (B) Sanger sequencing of the PCR amplified DNA from A mapped to the repetitive Stop cassette region in the Ai9-tdTomato mouse. The black lines represent DNA that is present and the white lines represent regions of deletions within the DNA.

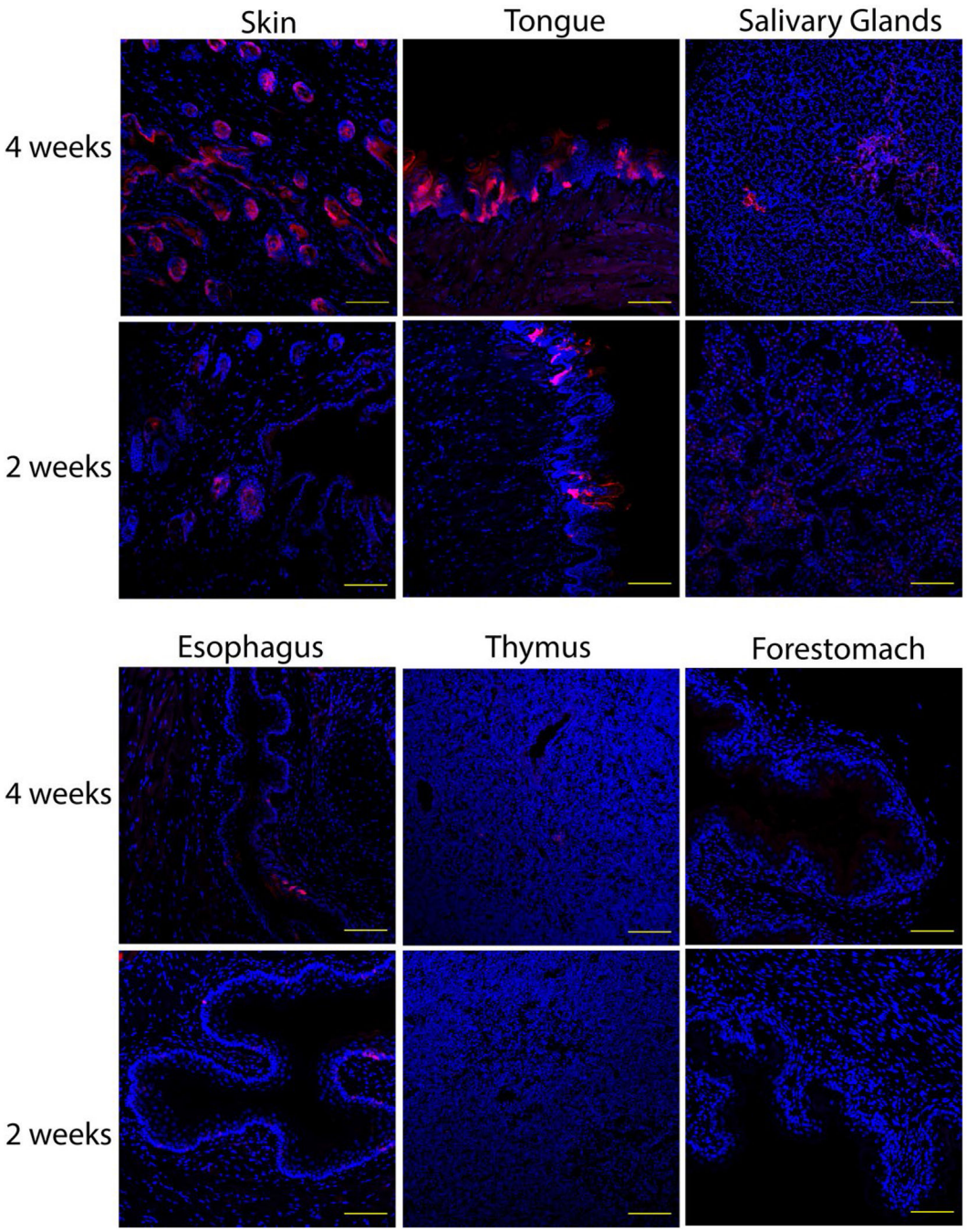

Figure 4: Doxycycline inducible Krt14-scgRNA.

(A) Krt14-scgRNA mice were bred to constitutively express Ai9-tdTomato, the reverse tetracycline response activator (rTTA) and the doxycycline inducible (TRE) promoter driving the expression of spCas9. In the presence of doxycycline, the rTTA can activate the TRE promoter to express spCas9. Representative sections taken from the tongue and skin of mice that had been fed a doxycycline diet for either 4 weeks (n=4) or 2 weeks (n=6). Scale bars represent 100 μm.

Table 1.

DNA sequences of the scgRNA transgene and single stranded repair template.

	Sequence
K14-scgRNA transgene	5’-AGGTTTCTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTCAAACCTCTACA AATGTGGTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATC AACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTGGCCGGCATGGTCCCAGCCTC CTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCGAATGGGAC-3’
Single stranded repair template	5’-GTGGTCTCCACCCACGAGCAGGTCCTGCGCACCAAGAACTAAAGCTGCTACATG CTGCTCAGGCTTAGGAGGCTCCTGCGTGGACGCAGATACTGCAGGTTTCTGATG AGTCCGTGAGGACGAAACGAGTAAGCTCGTCAAACCTCTACAAATGTGGTAGTTT TAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGT GGCACCGAGTCGGTGCTTTTGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCC GGCTGGGCAACATGCTTCGGCATGGCGAATGGGACTGGAAGAGCCCCTGTATT GTCCCTATAGGCTTCACTCTTTACTTGACCCTTGTCCTTCCTGGCAAGCAATAAA GCTTCTTTTTCTGAGTTGCACAA-3’