

Supplementary Figures

Genomic tagging of endogenous human ESCRT-I complex preserves ESCRT-mediated membrane remodeling functions

Huxley K. Hoffman¹, Melissa V. Fernandez², Nicholas S. Groves¹, Eric O. Freed², Schuyler B. van Engelenburg^{1*}

¹ Molecular and Cellular Biophysics Program, Department of Biological Sciences, University of Denver, Denver, CO 80210, USA

² Virus-Cell Interaction Section, HIV Dynamics and Replication Program, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA

(A) Tsg101 sgRNA target site

M A V S E S Q L K K M V *S K	Tsg101 exon 1
cggccgtcatggcgggtgtcggagagccagctcaagaaaatggtgt*ccaagtgaggctg	genomic DNA
<u>GCTCAAGAAAATGGTGT*CCAAGG</u>	sgRNA

(B) GFP-Tsg101 HDR template sequence

5' homology region, PuroR, P2A, EGFP, 3' homology region (with silent mutations)

Gggtgggtacagaggagaaattctgacttacggaatgatttctcggagaagaattactatctggcttcttgtgaaacaaagctctacacttta
 cccttgccaacgtatagatgaaggtctttaagcttaaaaaaatcattaatagaggaccactgggcatcatctaaggcaagtgtatgacttta
 gagtcgggagaaatttgcgcctgagctagaaatgcactagcctaccattatcacaagagcaagtcagtgacagaaacagttctctgaccgccagtg
 cagtgccctcccacttgatcaatcagtcattaaggctcactggcaagcccttcgacttcttccctctccggcactagcttccctaaagtcgactc
 tgtgaaatacctgctaattcttactcgaaccagagtttgcacctcggaaacagtccttgcaccatcttgggaccgattcgtcctgctgattcc
 ggcaactcgagggggcgggttgggtggaacccacccaccaggccctctcaatcccacacgggtgtccacccttgcccgcaggacggcggcccgaa
 gtgacgtagtggaagcgggaagtggtgtagtggtgcccacttctggtgttggaggccgggttgggggtggtgcgattgtgtgggacggctcggggc
 agcccagcagcggctgaccctctgcctcgggggaaggagtcgcccaggcggcggctcatgaccgagtaacaagcccaggtgcccctcggccaccg
 gacgacgtcccagggcgtacgcaccctcgcgcgcggttcgcccactaccocgccacgcgccacaccgctcgatccggaccgccacatcgagcg
 ggtcaccgagctgcaagaactcttctcaccgcgctcgggctcgacatcggcaaggtgtgggtcgggacgacggcggcggcggctggcggctgga
 ccacgcccggagagcgtcgaagcggggcggtgttcgcccagatcggcccgcgcatggccgagttgagcgggtcccggtggccgagcaacag
 atggaaggcctcctggcggcaccggcccaggagcccgcgtggttctgcccaccgctcggagctctcggccgaccaccagggcaagggctcggg
 cagcggcgtcgtgctcccggagtgaggcggcggagcgcgcgggggtgccgccttctggagacctcggcggccgcaacctccccttctacg
 agcggctcggctcaccgctcaccgcccagctcgaggtgccgaaggacggcgcacctggtgcatgaccgcaagcccgggtgcccgcagcggcgcc
 accaacttcagcctgctgaagcaggccggcgagctggaggagaaccccggccccatggtgagcaaggcggagagctgttcaccgggggtggtgcc
 catcctggtcgagctggacggcagctaaacggccacaagttcagcgtgtccggcgaggcggagggcgatgccacctacggcaagctgacctga
 agttcatctgcaccaccggcaagctgcccgtgcccgtgcccaccctcgtgaccacctgacctacggcgtgcaagtgcttcagccgctaccccgac
 cacatgaagcagcagcacttctcaagtccgcatgcccgaaggctacgtccaggagcgcaccatcttctcaaggacgacggcaactacaagac
 ccgcccggaggtgaagttcagggcgacacctgggtaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatcctggggcaca
 agctggagtacaactacaacagccacaacgtctatatcatggccgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatc
 gaggacggcagcgtgacgtcgcggaccactaccagcagaacaccccctcggcgacggcccccgtgctgctgcccgaaccactacctgagcac
 ccagtcgccctgagcaaaagaccccacgagaagcggcatcacatggtcctgctggagttcgtgaccgcccgggatcactctcggcatggacg
 agctgtacaagtcgggaatggcgggtgctggagagccagctcaa~~aa~~gatggt~~ca~~gcaaggtgaggctgacgacgctcgcctcccagggcgcgc
 caccgctccctcggcgcctgtcagttccgtcccggcccagccaagcaagcttcccagacggcgggaagcccgggtgcagtccttagcgaact
 cctcagaaccccggcccggcctgtcgcctggtcgggaatcccctacgggagctgggaggggtgggggacggcgacagtcaacaaaggcgt
 ggagcggaggtacctgacacctgcccaccocgcctcctctcttccactgagtttggagctgtctgtggggcagtgtagttttctggttgtt
 ttataaacacaaaggatcatacttagttgtagatctgaggcaatcctctactcctgtcccaatctagggatcctccagctcttagccttct
 tttgcccagcccagagagtggtggcctttcagttttgtattctccagaaaagaagcgaataaccagtccttactgcccaggtgaaataca
 taccattaccatttcagggttttctaagagtaactacttttaagtgtagcttcttataacgcgtatttatgttggcttgagcttaattatgat
 cttcacggctcaag

Figure S1. Guide RNA and HDR donor template sequences for GFP-Tsg101 knock-in. (A) Tsg101 sgRNA and its target site in Tsg101 Exon 1. The protospacer sequence is underlined and the protospacer adjacent motif (PAM) is *italicized*. The Cas9 nuclease cuts the DNA 3-4 nucleotides upstream of the PAM (*). (B) Annotated sequence of the GFP-Tsg101 HDR template, containing a PuroR-P2A-EGFP cassette flanked by homology regions matching the genomic DNA on either side of the target site.

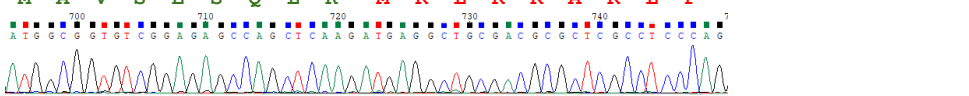
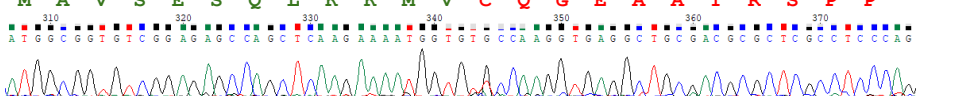
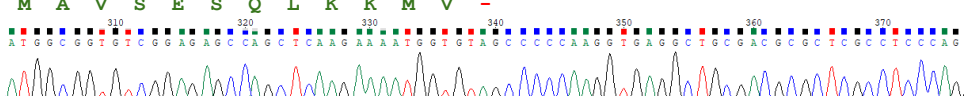
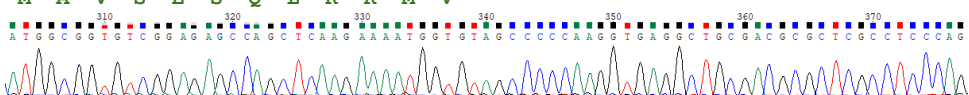
WT reference sequence	<p>ATGGCGGTGTCGGAGAGCCAGCTCAAGAAAATGGTGT*CCAAGGTGAGGCTGCGACGCGCTCGCTCCCAG</p> <p>M A V S E S Q L K K M V *S K</p>	<p>(*) sgRNA target site. Translation: Tsg101 Exon 1.</p>
KI HeLa A	<p>ATGGCGGTGTCGGAGAGCCAGCTCAAGA*TGAGGCTGCGACGCGCTCGCTCCCAG</p> <p>M A V S E S Q L K M R L R R A R L P</p> 	<p>* 15 nt deletion removes splice site.</p>
KI HeLa B	<p>ATGGCGGTGTCGGAGAGCCAGCTCAAGAAAATGGTGTGCCAAGGTGAGGCTGCGACGCGCTCGCTCCCAG</p> <p>M A V S E S Q L K K M V C Q G E A A T R S P P</p> 	<p>1 nt insertion causes frameshift.</p>
KI Jurkat A	<p>ATGGCGGTGTCGGAGAGCCAGCTCAAGAAAATGGTGTAGCCTCAAGGTGAGGCTGCGACGCGCTCGCTCCCAG</p> <p>M A V S E S Q L K K M V -</p> 	<p>5 nt insertion introduces premature stop codon.</p>
KI Jurkat B	<p>ATGGCGGTGTCGGAGAGCCAGCTCAAGAAAATGGTGTAGCCTCAAGGTGAGGCTGCGACGCGCTCGCTCCCAG</p> <p>M A V S E S Q L K K M V -</p> 	<p>5 nt insertion introduces premature stop codon.</p>

Figure S2. GFP-Tsg101 knock-in cell lines have knockout indels on their untagged Tsg101 alleles. To sequence the non-knock-in alleles of Tsg101, PCR was performed on genomic DNA of the knock-in cell lines with primers that anneal to the genomic DNA on either side of the Tsg101 target site. The primers could also anneal to a knock-in allele, but that would give a much larger PCR amplicon, which was excluded by limiting the extension time and by gel-purifying the band at the predicted size for a non-knock-in allele. Sanger sequencing of the PCR products showed indels that knock-out expression of Tsg101. The indels, and their effects on the translated protein sequence, are indicated in red. This confirms that the knock-in cell lines have no functional alleles for untagged Tsg101.

(A) 5' knock-in junction

HumanGenomeRef	1	GATTCGTCCTGCTGATTCCGGCACTCGAGGGCGGGTGGTTGGAACCCACCC	CCAGGCCCTCTCAATCCCACACGG
DonorPlasmid	1	GATTCGTCCTGCTGATTCCGGCACTCGAGGGCGGGTGGTTGGAACCCACCC	CCAGGCCCTCTCAATCCCACACGG
WT_HeLa	1	GATTCGTCCTGCTGATTCCGGCACTCGAGGGCGGGTGGTTGGAACCCACCCGCCAGGCCCTCTCAATCCCACACGG	
KI_HeLa_A	1	GATTCGTCCTGCTGATTCCGGCACTCGAGGGCGGGTGGTTGGAACCCACCCGCCAGGCCCTCTCAATCCCACACGG	
KI_HeLa_B	1	GATTCGTCCTGCTGATTCCGGCACTCGAGGGCGGGTGGTTGGAACCCACCCGCCAGGCCCTCTCAATCCCACACGG	
WT_Jurkat	1	GATTCGTCCTGCTGATTCCGGCACTCGAGGGCGGGTGGTTGGAACCCACCCGCCAGGCCCTCTCAATCCCACACGG	
KI_Jurkat_A	1	GATTCGTCCTGCTGATTCCGGCACTCGAGGGCGGGTGGTTGGAACCCACCCGCCAGGCCCTCTCAATCCCACACGG	
KI_Jurkat_B	1	GATTCGTCCTGCTGATTCCGGCACTCGAGGGCGGGTGGTTGGAACCCACCCGCCAGGCCCTCTCAATCCCACACGG	
HumanGenomeRef	81	TGTCACCCCTTGCCCGCAGGACGGCGGCCCGGAAGTACGTAGTGAAGCGGAAGTGGTGTAGTGGTGCCGACTTCCTGT	
DonorPlasmid	81	TGTCACCCCTTGCCCGCAGGACGGCGGCCCGGAAGTACGTAGTGAAGCGGAAGTGGTGTAGTGGTGCCGACTTCCTGT	
WT_HeLa	81	TGTCACCCCTTGCCCGCAGGACGGCGGCCCGGAAGTACGTAGTGAAGCGGAAGTGGTGTAGTGGTGCCGACTTCCTGT	
KI_HeLa_A	81	TGTCACCCCTTGCCCGCAGGACGGCGGCCCGGAAGTACGTAGTGAAGCGGAAGTGGTGTAGTGGTGCCGACTTCCTGT	
KI_HeLa_B	81	TGTCACCCCTTGCCCGCAGGACGGCGGCCCGGAAGTACGTAGTGAAGCGGAAGTGGTGTAGTGGTGCCGACTTCCTGT	
WT_Jurkat	81	TGTCACCCCTTGCCCGCAGGACGGCGGCCCGGAAGTACGTAGTGAAGCGGAAGTGGTGTAGTGGTGCCGACTTCCTGT	
KI_Jurkat_A	81	TGTCACCCCTTGCCCGCAGGACGGCGGCCCGGAAGTACGTAGTGAAGCGGAAGTGGTGTAGTGGTGCCGACTTCCTGT	
KI_Jurkat_B	81	TGTCACCCCTTGCCCGCAGGACGGCGGCCCGGAAGTACGTAGTGAAGCGGAAGTGGTGTAGTGGTGCCGACTTCCTGT	
HumanGenomeRef	161	TGTTTGAGGCCGGGTGGGGTGTGCGATTGTGTGGACGGTCTGGGCAGCCAGCAGCGGCTGACCOCTTGCCCTGCGG	
DonorPlasmid	161	TGTTTGAGGCCGGGTGGGGTGTGCGATTGTGTGGACGGTCTGGGCAGCCAGCAGCGGCTGACCOCTTGCCCTGCGG	
WT_HeLa	161	TGTTTGAGGCCGGGTGGGGTGTGCGATTGTGTGGACGGTCTGGGCAGCCAGCAGCGGCTGACCOCTTGCCCTGCGG	
KI_HeLa_A	161	TGTTTGAGGCCGGGTGGGGTGTGCGATTGTGTGGACGGTCTGGGCAGCCAGCAGCGGCTGACCOCTTGCCCTGCGG	
KI_HeLa_B	161	TGTTTGAGGCCGGGTGGGGTGTGCGATTGTGTGGACGGTCTGGGCAGCCAGCAGCGGCTGACCOCTTGCCCTGCGG	
WT_Jurkat	161	TGTTTGAGGCCGGGTGGGGTGTGCGATTGTGTGGACGGTCTGGGCAGCCAGCAGCGGCTGACCOCTTGCCCTGCGG	
KI_Jurkat_A	161	TGTTTGAGGCCGGGTGGGGTGTGCGATTGTGTGGACGGTCTGGGCAGCCAGCAGCGGCTGACCOCTTGCCCTGCGG	
KI_Jurkat_B	161	TGTTTGAGGCCGGGTGGGGTGTGCGATTGTGTGGACGGTCTGGGCAGCCAGCAGCGGCTGACCOCTTGCCCTGCGG	
HumanGenomeRef	241	GGAAGGGAGTCGCCAGGCGGCCGTCATG	-----
DonorPlasmid	241	GGAAGGGAGTCGCCAGGCGGCCGTCATGACCCGAGTACAAGCCC	-----
WT_HeLa	241	GGAAGGGAGTCGCCAGGCGGCCGTCATG	-----
KI_HeLa_A	241	GGAAGGGAGTCGCCAGGCGGCCGTCATGACCCGAGTACAAGCCC	-----
KI_HeLa_B	241	GGAAGGGAGTCGCCAGGCGGCCGTCATGACCCGAGTACAAGCCC	-----
WT_Jurkat	241	GGAAGGGAGTCGCCAGGCGGCCGTCATG	-----
KI_Jurkat_A	241	GGAAGGGAGTCGCCAGGCGGCCGTCATGACCCGAGTACAAGCCC	-----
KI_Jurkat_B	241	GGAAGGGAGTCGCCAGGCGGCCGTCATGACCCGAGTACAAGCCC	-----

(B) 3' knock-in junction

HumanGenomeRef	1	-----	ATGGCGGTGTCGGAGAGCCAGCTCAA	AA	ATGGTGTCCAA
DonorPlasmid	1	GGGATCACCTCTCGGCATGGACGAGCTGTACAAGTCGGGA	AATGGCGGTGTCGGAGAGCCAGCTCAA	AAA	AAGATGGTGTCCAA
WT_HeLa	1	-----	ATGGCGGTGTCGGAGAGCCAGCTCAA	AA	ATGGTGTCCAA
KI_HeLa_A	1	GGGATCACCTCTCGGCATGGACGAGCTGTACAAGTCGGGA	AATGGCGGTGTCGGAGAGCCAGCTCAA	AAA	AAGATGGTGTCCAA
KI_HeLa_B	1	GGGATCACCTCTCGGCATGGACGAGCTGTACAAGTCGGGA	AATGGCGGTGTCGGAGAGCCAGCTCAA	AAA	AAGATGGTGTCCAA
WT_Jurkat	1	-----	ATGGCGGTGTCGGAGAGCCAGCTCAA	AA	ATGGTGTCCAA
KI_Jurkat_A	1	GGGATCACCTCTCGGCATGGACGAGCTGTACAAGTCGGGA	AATGGCGGTGTCGGAGAGCCAGCTCAA	AAA	AAGATGGTGTCCAA
KI_Jurkat_B	1	GGGATCACCTCTCGGCATGGACGAGCTGTACAAGTCGGGA	AATGGCGGTGTCGGAGAGCCAGCTCAA	AAA	AAGATGGTGTCCAA
HumanGenomeRef	42	GTGAGGCTGCGACGCGCTCGCCTCCCAGGGCGCGCCACCCTCCCTTCCGCGCCCTGTCGAGTCCGTCGCCGCCAGC			
DonorPlasmid	81	GTGAGGCTGCGACGCGCTCGCCTCCCAGGGCGCGCCACCCTCCCTTCCGCGCCCTGTCGAGTCCGTCGCCGCCAGC			
WT_HeLa	42	GTGAGGCTGCGACGCGCTCGCCTCCCAGGGCGCGCCACCCTCCCTTCCGCGCCCTGTCGAGTCCGTCGCCGCCAGC			
KI_HeLa_A	81	GTGAGGCTGCGACGCGCTCGCCTCCCAGGGCGCGCCACCCTCCCTTCCGCGCCCTGTCGAGTCCGTCGCCGCCAGC			
KI_HeLa_B	81	GTGAGGCTGCGACGCGCTCGCCTCCCAGGGCGCGCCACCCTCCCTTCCGCGCCCTGTCGAGTCCGTCGCCGCCAGC			
WT_Jurkat	42	GTGAGGCTGCGACGCGCTCGCCTCCCAGGGCGCGCCACCCTCCCTTCCGCGCCCTGTCGAGTCCGTCGCCGCCAGC			
KI_Jurkat_A	81	GTGAGGCTGCGACGCGCTCGCCTCCCAGGGCGCGCCACCCTCCCTTCCGCGCCCTGTCGAGTCCGTCGCCGCCAGC			
KI_Jurkat_B	81	GTGAGGCTGCGACGCGCTCGCCTCCCAGGGCGCGCCACCCTCCCTTCCGCGCCCTGTCGAGTCCGTCGCCGCCAGC			
HumanGenomeRef	122	CAAGCAAGCTTCCCAGACGGGCGGAAGCCCGGTGCAGTCTTAGCGACCTCCTCAGAACCCCGCCCGAGGGCGCTGT			
DonorPlasmid	161	CAAGCAAGCTTCCCAGACGGGCGGAAGCCCGGTGCAGTCTTAGCGACCTCCTCAGAACCCCGCCCGAGGGCGCTGT			
WT_HeLa	122	CAAGCAAGCTTCCCAGACGGGCGGAAGCCCGGTGCAGTCTTAGCGACCTCCTCAGAACCCCGCCCGAGGGCGCTGT			
KI_HeLa_A	161	CAAGCAAGCTTCCCAGACGGGCGGAAGCCCGGTGCAGTCTTAGCGACCTCCTCAGAACCCCGCCCGAGGGCGCTGT			
KI_HeLa_B	161	CAAGCAAGCTTCCCAGACGGGCGGAAGCCCGGTGCAGTCTTAGCGACCTCCTCAGAACCCCGCCCGAGGGCGCTGT			
WT_Jurkat	122	CAAGCAAGCTTCCCAGACGGGCGGAAGCCCGGTGCAGTCTTAGCGACCTCCTCAGAACCCCGCCCGAGGGCGCTGT			
KI_Jurkat_A	161	CAAGCAAGCTTCCCAGACGGGCGGAAGCCCGGTGCAGTCTTAGCGACCTCCTCAGAACCCCGCCCGAGGGCGCTGT			
KI_Jurkat_B	161	CAAGCAAGCTTCCCAGACGGGCGGAAGCCCGGTGCAGTCTTAGCGACCTCCTCAGAACCCCGCCCGAGGGCGCTGT			
HumanGenomeRef	202	CGCCTGGTGCGGGAATCCCCGTACGGGAGCTGGGAGGTTGGGGACGGCGACAGTCAACAAAGGCGTGGAGCGGAGGCTA			
DonorPlasmid	241	CGCCTGGTGCGGGAATCCCCGTACGGGAGCTGGGAGGTTGGGGACGGCGACAGTCAACAAAGGCGTGGAGCGGAGGCTA			
WT_HeLa	202	CGCCTGGTGCGGGAATCCCCGTACGGGAGCTGGGAGGTTGGGGACGGCGACAGTCAACAAAGGCGTGGAGCGGAGGCTA			
KI_HeLa_A	241	CGCCTGGTGCGGGAATCCCCGTACGGGAGCTGGGAGGTTGGGGACGGCGACAGTCAACAAAGGCGTGGAGCGGAGGCTA			
KI_HeLa_B	241	CGCCTGGTGCGGGAATCCCCGTACGGGAGCTGGGAGGTTGGGGACGGCGACAGTCAACAAAGGCGTGGAGCGGAGGCTA			
WT_Jurkat	202	CGCCTGGTGCGGGAATCCCCGTACGGGAGCTGGGAGGTTGGGGACGGCGACAGTCAACAAAGGCGTGGAGCGGAGGCTA			
KI_Jurkat_A	241	CGCCTGGTGCGGGAATCCCCGTACGGGAGCTGGGAGGTTGGGGACGGCGACAGTCAACAAAGGCGTGGAGCGGAGGCTA			
KI_Jurkat_B	241	CGCCTGGTGCGGGAATCCCCGTACGGGAGCTGGGAGGTTGGGGACGGCGACAGTCAACAAAGGCGTGGAGCGGAGGCTA			
HumanGenomeRef	282	CCTGACACCTGCGGCCACCCGCCCTCCTCTCTTCCACTGAGTTTGGAGCTGTCTGTGGGG			
DonorPlasmid	321	CCTGACACCTGCGGCCACCCGCCCTCCTCTCTTCCACTGAGTTTGGAGCTGTCTGTGGGG			
WT_HeLa	282	CCTGACACCTGCGGCCACCCGCCCTCCTCTCTTCCACTGAGTTTGGAGCTGTCTGTGGGG			
KI_HeLa_A	321	CCTGACACCTGCGGCCACCCGCCCTCCTCTCTTCCACTGAGTTTGGAGCTGTCTGTGGGG			
KI_HeLa_B	321	CCTGACACCTGCGGCCACCCGCCCTCCTCTCTTCCACTGAGTTTGGAGCTGTCTGTGGGG			
WT_Jurkat	282	CCTGACACCTGCGGCCACCCGCCCTCCTCTCTTCCACTGAGTTTGGAGCTGTCTGTGGGG			
KI_Jurkat_A	321	CCTGACACCTGCGGCCACCCGCCCTCCTCTCTTCCACTGAGTTTGGAGCTGTCTGTGGGG			
KI_Jurkat_B	321	CCTGACACCTGCGGCCACCCGCCCTCCTCTCTTCCACTGAGTTTGGAGCTGTCTGTGGGG			

Figure S3. GFP-Tsg101 knock-in alleles show precise insertion. (A) 5' and (B) 3' homology junctions of the KI alleles were amplified by PCR from genomic DNA, analyzed by Sanger sequencing, and compared to the human genome reference sequence, donor plasmid sequence, and sequences from the parental (WT) cell lines. Non-coding regions (5'UTR and intron) are shown in black/white text, **Tsg101 exon 1 in blue**, **PuroR in red**, and **GFP in green**. The results show precise integration of the PuroR-P2A-GFP insert, with no abnormalities except for several single-nucleotide polymorphisms in non-coding regions, some of which were also found in the parental lines. The discrepancies in the Tsg101 coding sequence in the alignment are the silent mutations intentionally introduced to the donor plasmid to prevent CRISPR from cutting the knock-in allele.

