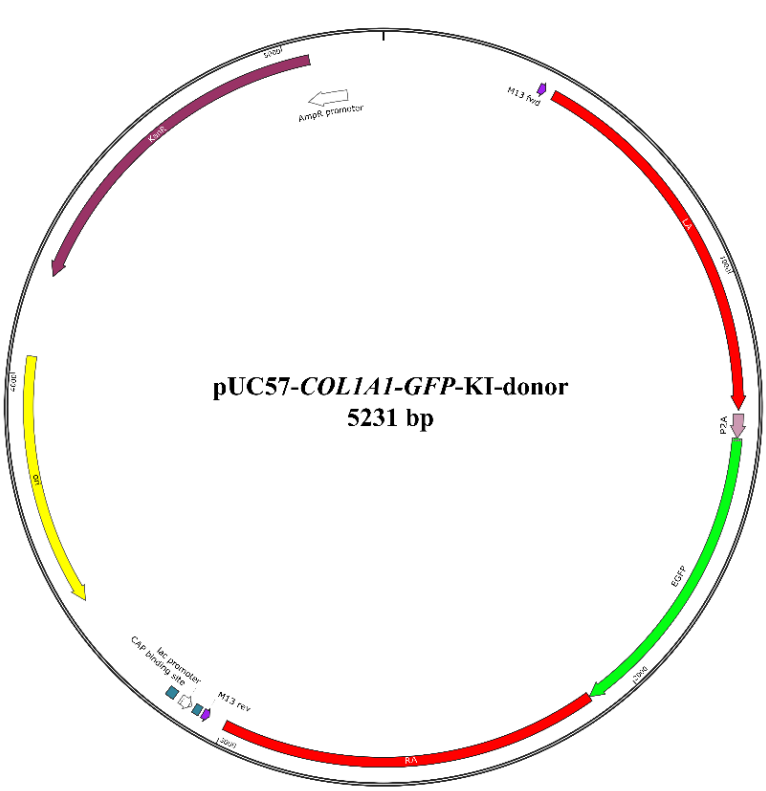
Appendix A Primers used in this study.

Primers	sequence（5'-3'）	Amplicon
COL1A1-sgRNA-F	CACCGAAACTCCCTCCGCCCCAATC	
COL1A1-sgRNA-R	AAACGATTGGGGCGGAGGGAGTTTC	
COL1A1 EX51F	ACACCCCTCTCCCATTGTCT	247 bp
COL1A1 EX51R	TGTTTGTTTCCAGGGTCAGG	
P1	CCTCCTCCTCAGCCACATAG	1271 bp
P2	GAACTTCAGGGTCAGCTTGC	
P3	CGACCACTACCAGCAGAACA	1325 bp
P4	GGGAAGAAGAAAGGGTGGAG	
Prediated-OFF-Target1-F	TGGACCACGTTCCCTTTAAC	507 bp
Prediated-OFF-Target1-R	TAAGGGCAGGTGAGGAAATG	
Prediated-OFF-Target2-F	CGCTCTAAAACCCAGACCAG	581 bp
Prediated-OFF-Target2-R	CAAGGGGATGTGCTTAGGAA	
Prediated-OFF-Target3-F	TGAAGCTCTCACGTCTCCAA	527 bp
Prediated-OFF-Target3-R	ACAAATTTTCCCTGCACGAC	
Prediated-OFF-Target4-F	TGATGCAGGCCACACTTTAG	629 bp
Prediated-OFF-Target4-R	GTATGGGACAGCGCTTGTTT	
Prediated-OFF-Target5-F	TGGGGGTGTGAATCCTCTAA	604 bp
Prediated-OFF-Target5-R	TTCAGCTCTGGGACACTCAA	
Prediated-OFF-Target6-F	GCACTCCCCTTTCCTCTTTT	535 bp
Prediated-OFF-Target6-R	TTCAGGCAAGGCTTTATTGG	
Prediated-OFF-Target7-F	GAGCTGGGAGATGGTCTTTG	598 bp
Prediated-OFF-Target7-R	CTAAGCCAGTGTCCCTTTGC	
Prediated-OFF-Target8-F	CAGGAGTGGGATTGCTGAGT	523 bp
Prediated-OFF-Target8-R	CATTGGGAGTTAGGGCTTCA	
Prediated-OFF-Target9-F	TGCTTTTCCTTAGCCCAAGA	567 bp
Prediated-OFF-Target9-R	AGACCTGTTTGGGGACCTCT	
COL1A1-125bp-qF	GGGTTCGGAGGAAAGTCAGG	125 bp
COL1A1-125bp-qR	TTAGGCAAGGAACAGAGCGG	
ACTB-138bp-qF	GGACTTCGAGCAGGAGATGG	138 bp
ACTB-138bp-qR	AGGAAGGAGGGCTGGAAGAG	

The bold bases represent stick end for ligation.

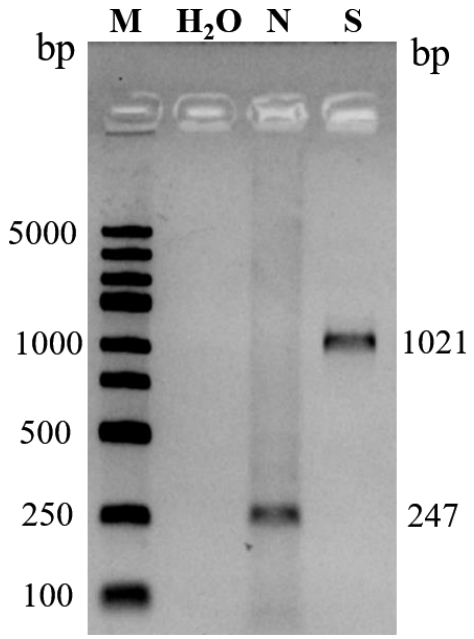


Appendix B Vector map of the donor vector pUC57-*COL1A1*-*GFP*-KI-donor. The *2A*-*GFP* fragment flanking homology arms were synthesized and inserted into pUC57.

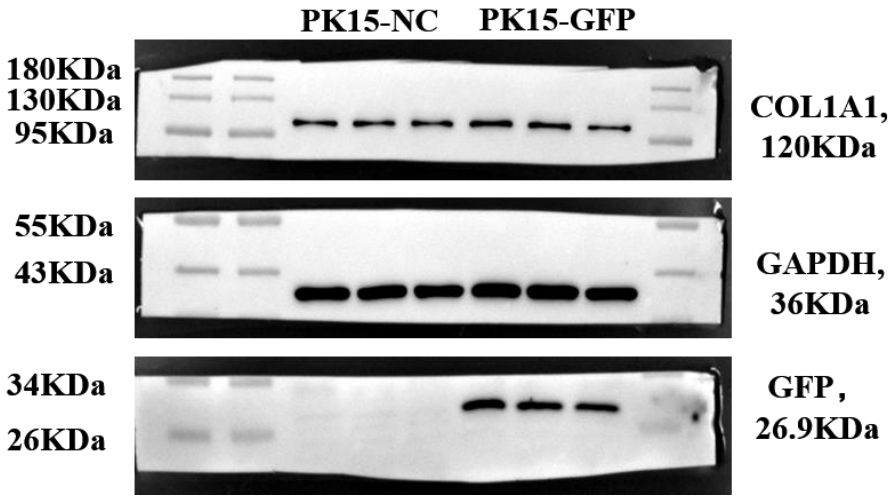
Appendix C Detailed donor vector sequences.The sequences of the homolog arms and *2A*-*GFP* fragments are shown in colors corresponding to those of the vector map mentioned in **Appendix B**.

tcgcgcgtttcggtgatgacggtgaaaacctctgacacatgcagctcccggagacggtcacagcttgtctgtaagcggatgccgggagcagacaagcccgtcagggcgcgtcag
cgggtgttggcgggtgtcggggctggcttaactatgcggcatcagagcagattgtactgagagtgcaccatatgcggtgtgaaataccgcacagatgcgtaaggagaaaataccgc
atcaggcgccattcgccattcaggctgcgcaactgttgggaagggcgatcggtgcgggcctcttcgctattacgccagctggcgaaaggggatgtgctgcaaggcgattaagttg
ggtaacgccagggttttcccagtcacgacgttgtaaaacgacggccagtgaattgacgcgtattgggatGACCTGCGTGTACCCCACTCAGCCCAGCGT
GCCCCAGAAGAACTGGTACATCAGCAAGAACCCCAAGGACAAGAGGCACGTCTGGTACGGCGAGAGCATGAC
CGACGGATTCCAGGTGCGTGAACTGGAGCCACTCTTCCAAGATGGGCTGGCCCAGGGCTCCACAGGGGGGTT
GATCCTGGGTAACCCAGACCTCTGCTACGTGTGACGGGCTGGGAGGCAGGACCCCGGAGTCCTCGAGACTGG
CTGGGACACAGGGTACCTGTAGGCAAGGGGTGCCCCTCTGGACGGGGGGTTCCCTGGGGTATCTTCAGTGGGA
CGTCAGGGAGGGAGGCCAGGGACCAGAGGCTCCGGCACGGTCGGTCCCTCACTCCTCGCTCTCCCACCCCCC
CCAGTTCGAGTACGGCGGCGAGGGCTCCGATCCTGCTGACGTGGCCATCCAGCTGACCTTCCTGCGCCTGATG
TCCACTGAGGCTTCCCAGAACATCACCTACCACTGCAAGAACAGCGTGGCCTACATGGACCAGCAGACTGGCA
ACCTCAAGAAGGCCCTGCTCCTCCAGGGCTCCAACGAGATCGAGATCCGGGCCGAGGGCAACAGCCGCTTCA
CCTACAGCGTGATCTACGACGGCTGCACGGTGAGTCCAGCAGGGCCCCCATTCATCGGGCTCTGGCTTCCCAG
GCAAGCCCAGTGCCCCCCACCCGCTCACTCCCGCTGGGTGACACCCCTCTCCCATTGTCTCCCCTTCACCCCAG
AGTCACACCGGAGCCTGGGGCAAGACAGTGATCGAATACAAAACCACCAAGACCTCCCGCCTGCCCATCATCG
ATGTGGCCCCCTTGGACGTTGGCGCCCCCGACCAAGAATTCGGCATCGACCTTAGCCCTGTCTGCTTCCTGGGA
AGCGGAGCCACCAACTTCAGCCTGCTGAAGCAGGCCGGCGACGTGGAGGAGAACCCCGGCCCCATGGTGAGC
AAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAG
TTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCG
GCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCC
CGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTC
TTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATC
GAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGC

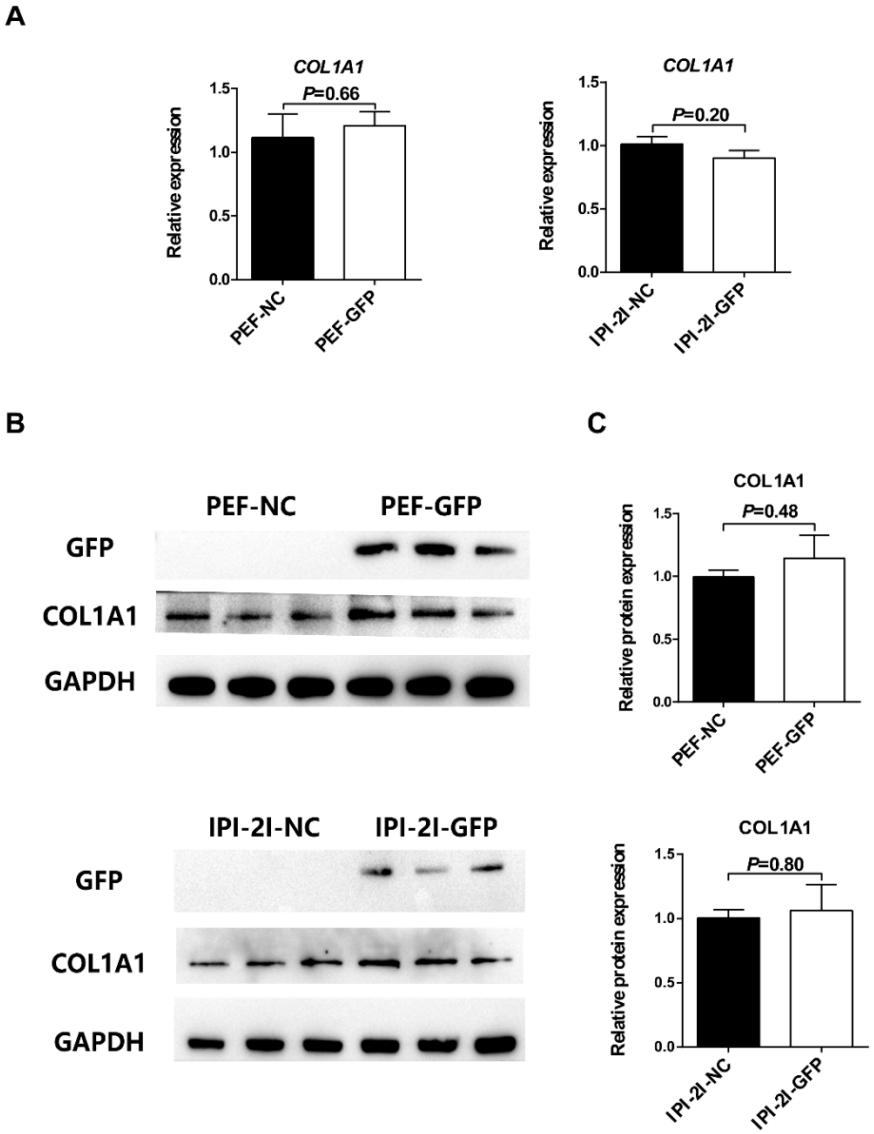
CACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCG
AGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGC
CCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCC
TGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGCCCCAATCTGGCT
CCCTCCCACCCAACCCAATTGCCCCTGACCCTGGAAACAAACAAACACCCCAAACGGAAACCCCCCCAAAAG
CCAAAAAATGGCAGACAATTTCACATGGACTTTGGAAAATATTTTTTTCCTTTGCATTCATCTCTCAAACTTAGT
TTTTAAATCTTTGACCAACTGAACATGACCAAAAACCAAAAGTGCATTTCAAACCTTACCAAAAAAAAAAAAAA
AAAAAAAGAATAAATAAATAACTTTTTAAAAAAGGAAGCTTGGTCCACTTGCTTGAAGACCCATGTGGGGGTA
AGTCCTTTTCTGCCCGTTGGGCTTATGACACCCCAACTCTGCCCTTTCTGCTCCTTTCTCCATGCCTTCCTGGGG
CCTCCCCTCCACTGCTCCCCAAATCTGAGTCTCCCCCAAAGACACAGAAAACAATGCATTGTCTGCCCAGCAA
CCAAAGGCAATGCTGAAACACCCAGTGGCCCCCCACCCCCAGCCCACCCCCCTCACCCAGCACCCCCAAACC
CTGGGGGCACCTGGAGTTCACGGACTGCCAAGGAAGCCTTTACCATCTGGCGTCCCCGTCGCCCTGGCAACAT
ACCCCTCTTTGTTTTGGAGGGGAGCATGCCAGGGAGACCACCGGCCCTTCACCACCGGGTTCGGAGGAAAGTC
AGGAGGGGCCAAGACAGAGCAGAAACCTCGGATTTGGGACCCAAACGGGCCAGAGCCCCGTGCAACCTGGC
TGGGTGGGAGAGACTGACCGCTCTGTTCCTTGCCTAATTGTGTTGCTGAAAGACTACCTCGTTCTTGTCTTTGT
GTGTCACCGGGGCAACTGCGTGGGGGCGGGGGTGGGGGCGGGGTGGCAGGGGCTCCTCATTTTATACCAAAG
GTGCTACCAAatcccaatggcgcgccgagcttggctcgagcctcgagcatggtcatagctgtttcctgtgtgaaattgttatccgctcacaattccacacaacatacgagccg
gaagcataaagtgtaaagcctggggtgcctaatgagtgagctaactcacattaattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatc
ggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggc
ggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccatagg
ctccgccccccctgacgagcatcacaaaatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggaagctccctcgt
gcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgct
ccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagcc
actggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaagaacagtatttggtatctgcgctctgctgaagc
cagtgttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaag
aagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaa
gttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttagaaaaactcatcgagcatcaaatgaaactgcaatttattcatatcaggattatcaataccatatttttgaaaaagc
cgtttctgtaatgaaggagaaaactcaccgaggcagttccataggatggcaagatcctggtatcggtctgcgattccgactcgtccaaccatcaatacaacctattaatttcccctcgtc
aaaaataaggttatcaagtgagaaatcaccatgagtgacgactgaatccggtgagaatggcaaaagtttatgcatttctttccagacttgttcaacaggccagccattacgctcgtcatc
aaaatcactcgcatcaaccaaaccgttattcattcgtgattgcgcctgagcgagacgaaatacgcgatcgctgttaaaaggacaattacaaacaggaatcgaatgcaaccggcgcag
gaacactgccagcgcatcaacaatattttcacctgaatcaggatattcttctaatacctggaatgctgttttcccagggatcgcagtggtgagtaaccatgcatcatcaggagtacggat
aaaatgcttgatggtcggaagaggcataaattccgtcagccagtttagtctgaccatctcaatctgtaacatcattggcaacgctacctttgccatgtttcagaaacaactctggcgcatc
gggcttcccatacaatcgatagattgtcgcacctgattgcccgacattatcgcgagcccatttatacccatataaatcagcatccatgttggaatttaatcgcggcctagagcaagacgtt
tcccgttgaatatggctcatactcttcctttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaatagggttccgcgc
acatttccccgaaaagtgccacctgacgtctaagaaaccattattatcatgacattaacctataaaaataggcgtatcacgaggccctttcgtc



Appendix D Genomic PCR analysis of PK15-GFP cells. The *2A-GFP* fragment was specificly inserted at both alleles of *COL1A1* gene in the PK15-GFP cells. M, DL5000 marker; H_2O, negative control; N, wild type PK15 cells; S, PK15-GFP cells.



Appendix E Western blotting for GFP and COL1A1 expression in PK15-GFP and wild type PK15 cells (n=3). PK15-NC, wild type PK15 cells; PK15-GFP, a homozygous *GFP* inserted PK15 cell clone.



Appendix F Expression of *COL1A1* gene in PEF and IPI-2I cells (n=3). (A) Transcription level of *COL1A1* gene in PEF and IPI-2I cells (n=3). *ACTB* was used as the reference gene for qPCR assay. PEF-NC, wild type PEF cells; PEF-GFP, GFP positive PEF cells, IPI-2I-NC, wild type IPI-2I cells, IPI-2I-GFP, GFP positive IPI-2I cells. (B) Western blotting for GFP and COL1A1 protein expression in PEF and IPI-2I cells (n=3). (C) Histogram showing the quantitative calculation of protein expression levels from Western blot images with ImageJ software (n=3). All data are presented as the means ± standard errors (SE), *P* values less than 0.05 were considered significant.

Appendix G Statistics of RNA-seq data

Groups	Samples	Raw Reads	Clean Reads	Clean Reads Rate(%)	Clean Bases	Clean Q30 Bases Rate(%)	Mapped Reads	Mapping Rate
NC	PEF_NC1	47,766,398	46,200,182	96.72	6,930,027,300	92.44	44,322,049	95.93
	PEF_NC2	43,917,956	42,728,076	97.29	6,409,211,400	91.90	40,931,984	95.80
	PEF_NC3	46,669,088	45,315,010	97.10	6,797,251,500	92.77	43,691,153	96.42
GFP	PEF_GFP1	47,625,538	45,923,374	96.43	6,888,506,100	92.66	44,013,831	95.84
	PEF_GFP2	40,822,212	39,660,682	97.16	5,949,102,300	92.74	38,026,745	95.88
	PEF_GFP3	46,352,670	44,934,410	96.94	6,740,161,500	92.76	43,060,178	95.83