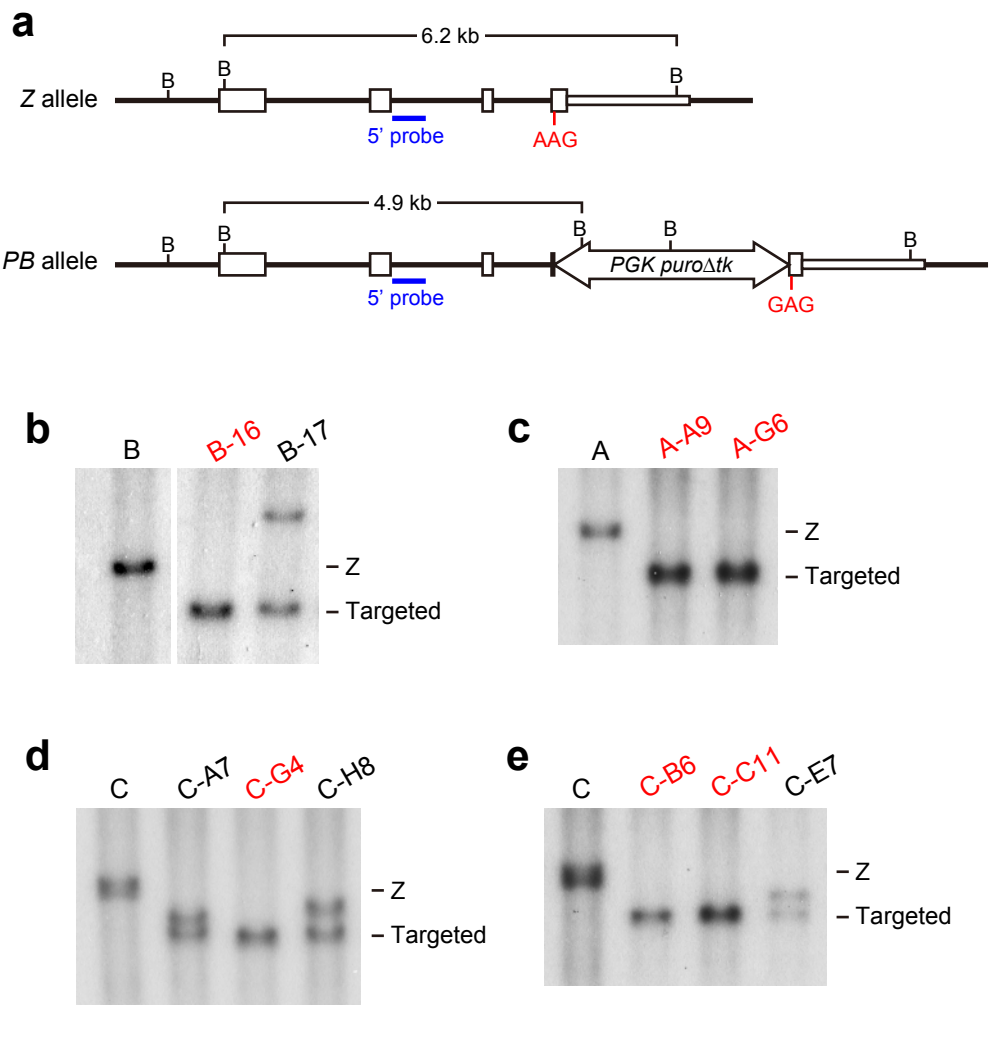
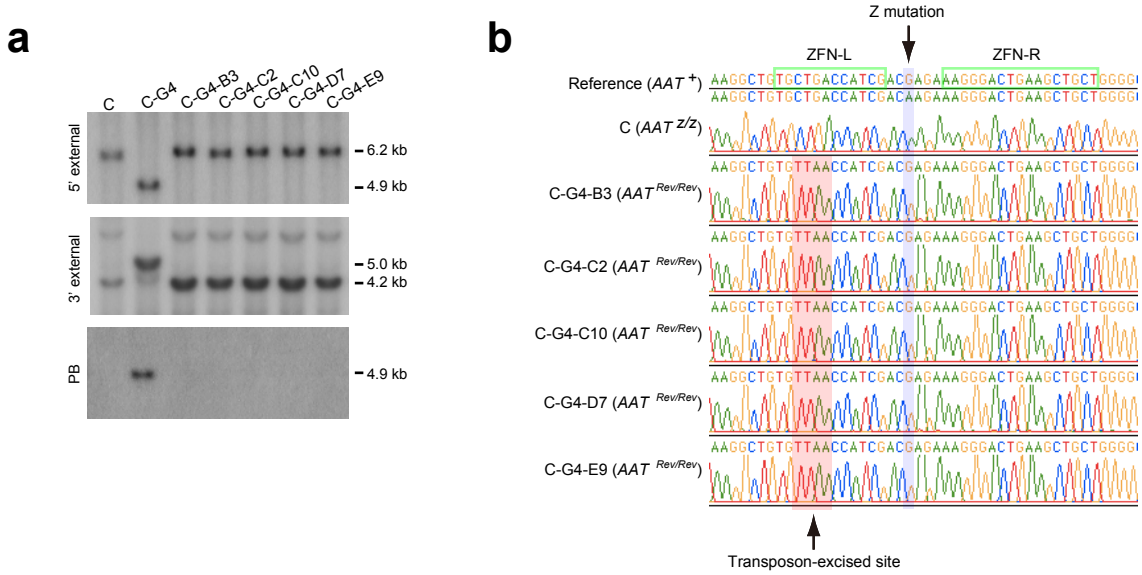
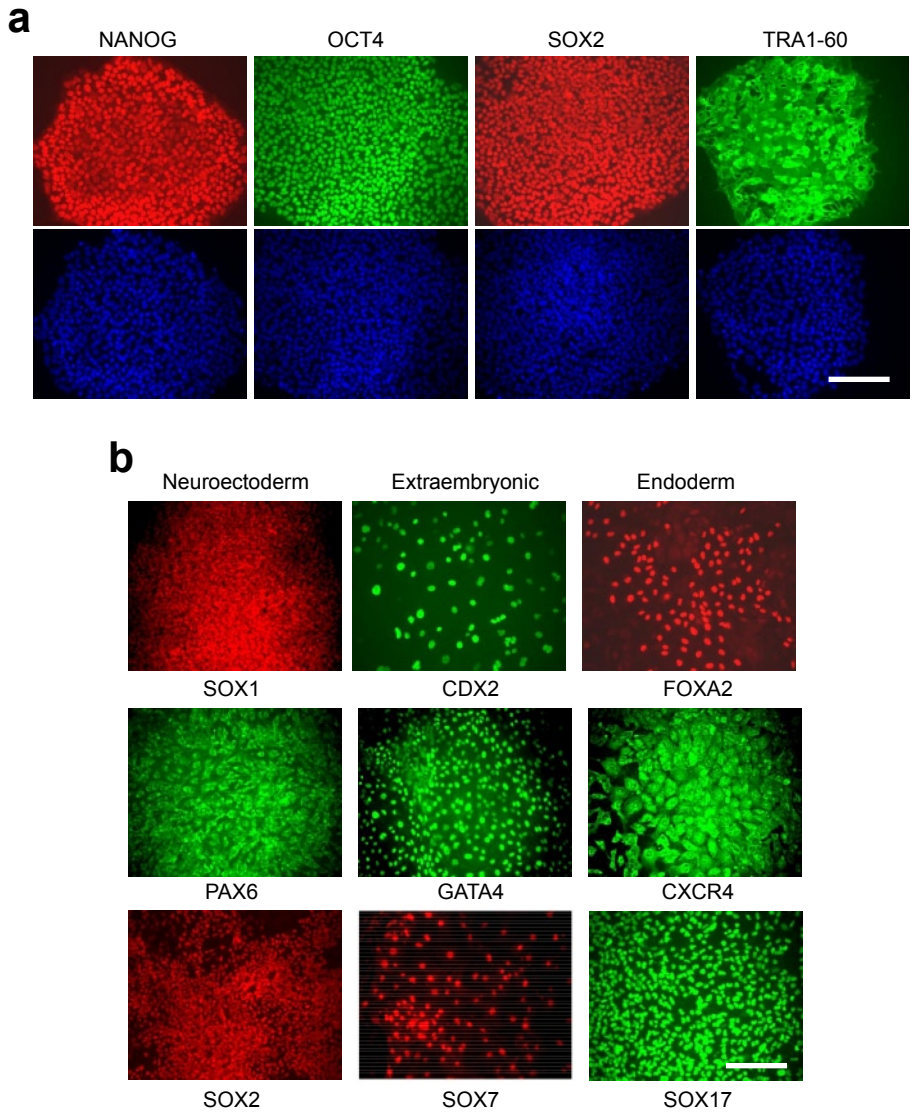


Supplementary Figure 1. Southern blot analyses of bi-allelic targeting of the *A1AT* gene in 3 independent A1ATD-hIPSC lines. **a**, The structures of the *Z* allele and the targeted *PB* allele. Thin and thick open boxes, non-coding and coding exon, respectively; B, BamHI. **b**, The result obtained with A1ATD-hIPSCs from patient B. Both alleles in line B-16 were correctly targeted while line B-17 carried one correctly targeted allele and one abnormal allele. **c**-**e**, Similar results were obtained with A1ATD-hIPSCs generated from patient A (**c**) and from patient C (**d**, **e**). Genomic DNA was digested by BamHI and hybridized with the 5' external probe. The first lane on each panel is DNA from the parental A1ATD-hIPSC line. Correct homozygous clones are highlighted in red.



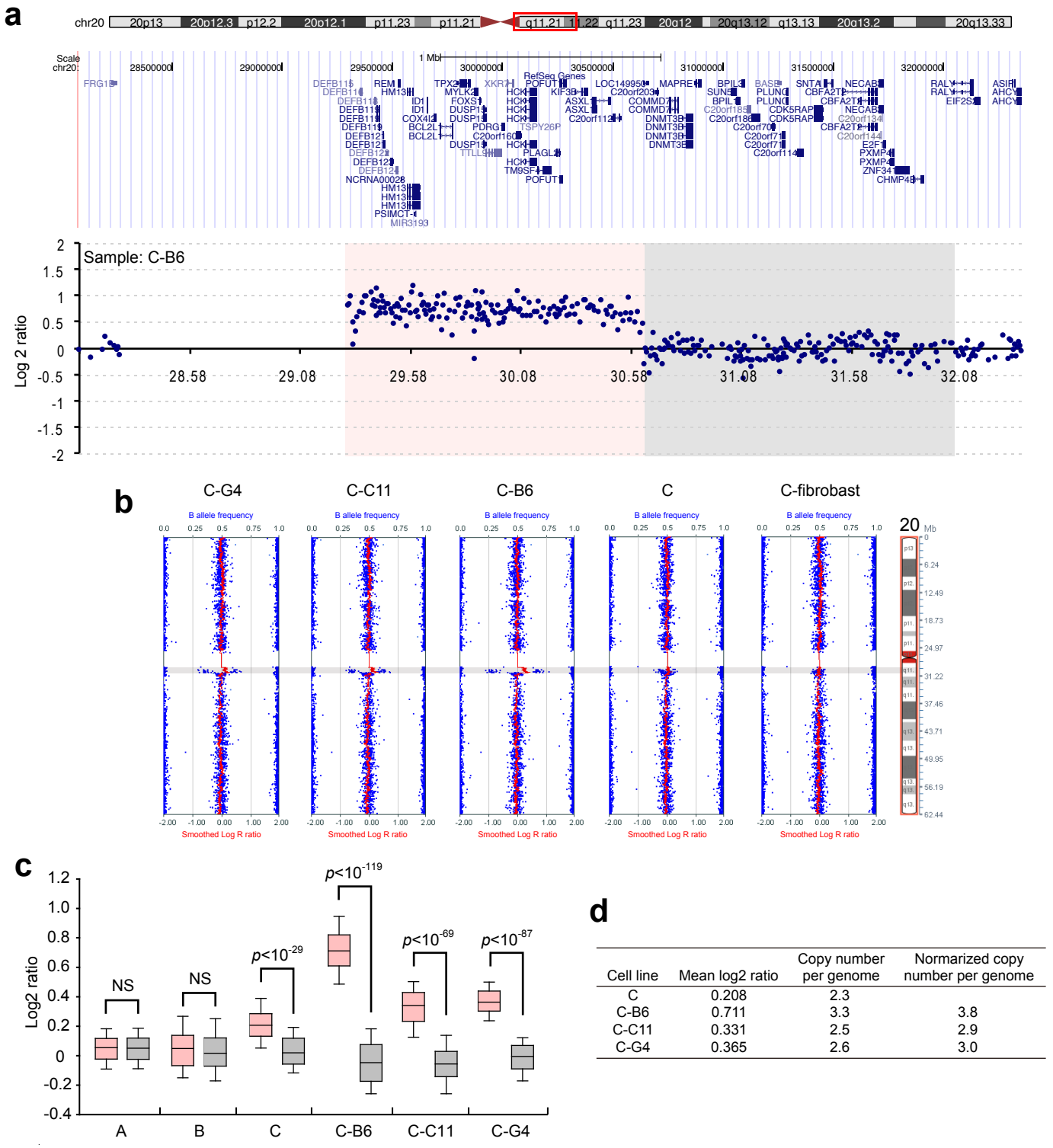
Supplementary Figure 2. Characterization of genome modification in A1ATD-hIPSC line C.

a, **b**, Southern blot (**a**) and sequence (**b**) analyses showing bi-allelic correction of Z mutation in A1ATD-hIPSCs generated from patient C. See Fig. 2 for details.



Supplementary Figure 3. Pluripotency of corrected hIPSCs.

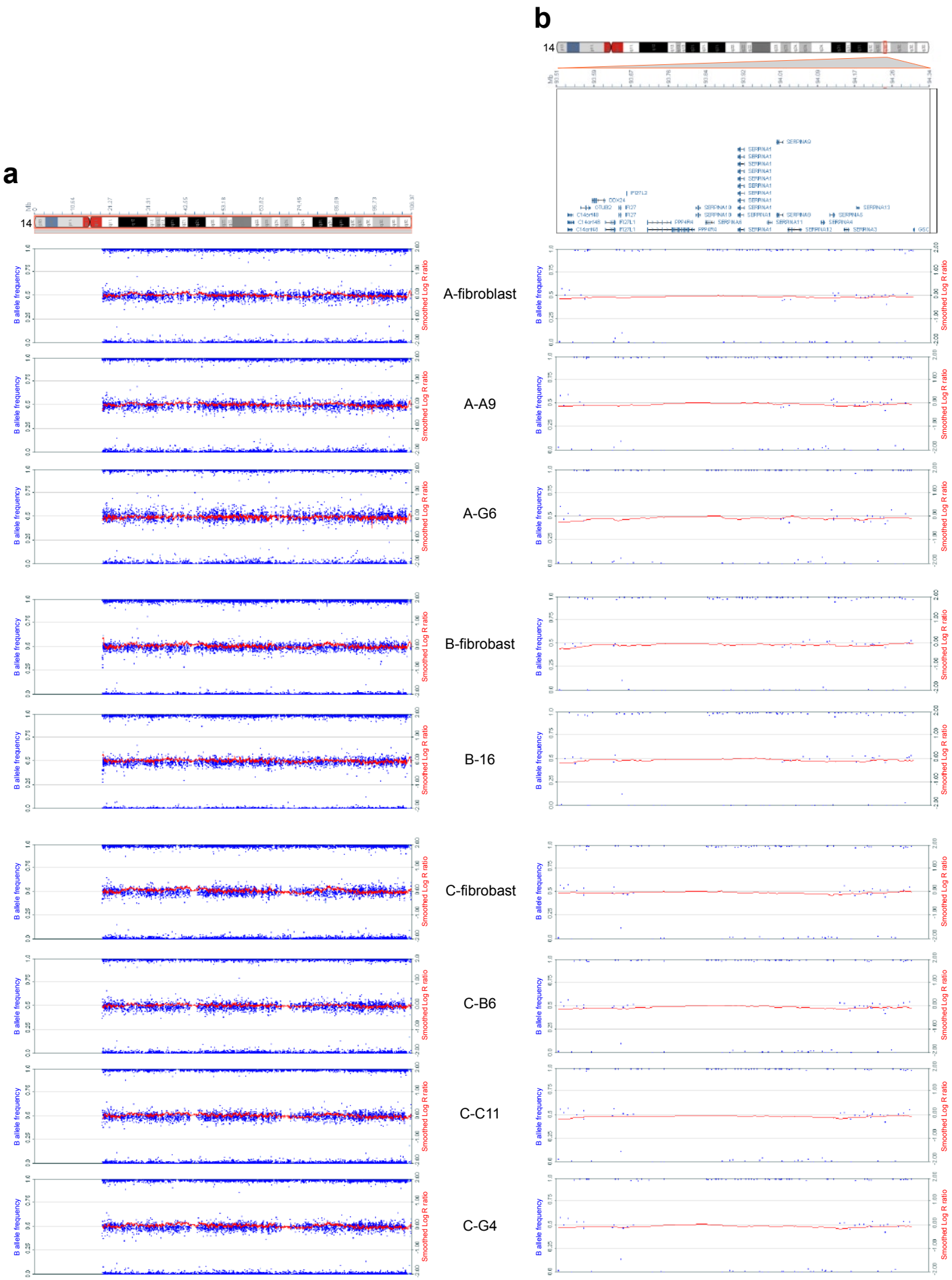
a, Immunofluorescence showing the expression of pluripotency markers OCT4, SOX2, NANOG and TRA1-60 in c-hIPSCs grown for 20 passages after correction. DNAs were stained with DAPI (lower panels). **b**, Immunofluorescence of *in vitro* differentiated c-hIPSCs showing maintenance of differentiation capacity into three germ layers. Scale bars, 200 μm



Cell line	Mean log2 ratio	Copy number per genome	Normarized copy number per genome
C	0.208	2.3	
C-B6	0.711	3.3	3.8
C-C11	0.331	2.5	2.9
C-G4	0.365	2.6	3.0

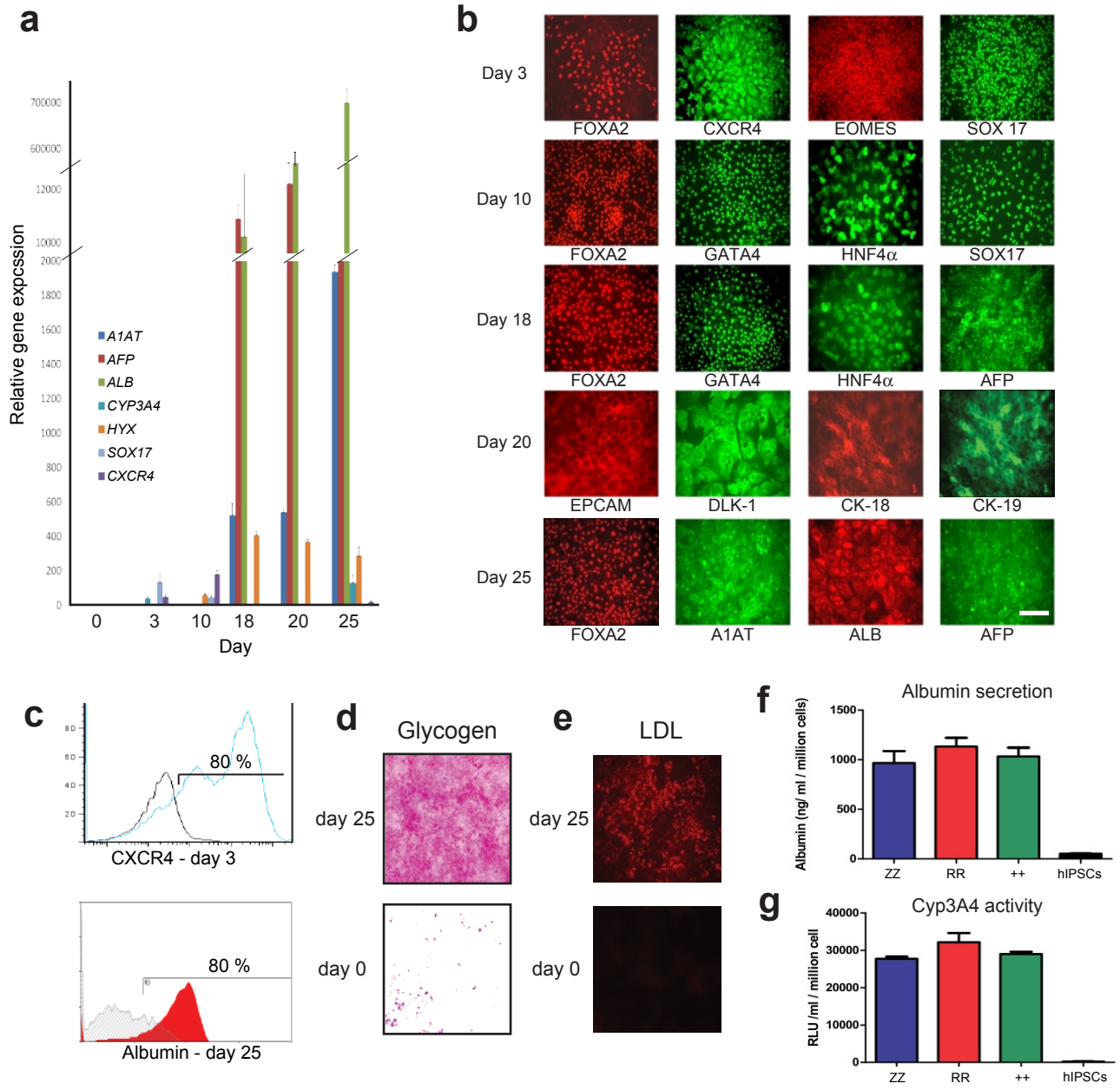
Supplementary Figure 4. Detailed analysis of 20q11.21 amplification

a, aCGH result of the homozygous clone, C-B6. Top, an ideogram of chromosome 20; middle, genes in the region pointed by red in the chromosome ideogram; bottom, the log2 ratio plot of C-B6. The amplified region is highlighted in red. The log2 ratios in the adjacent region highlighted in gray are used as a normal copy number control for a comparison shown in **c**. **b**, SNP data showing B allele frequency (*x* axis on top) and smoothed log *R* ratio (*x* axis on bottom) of chromosome 20 of indicated cell lines. The 20q11.21 is highlighted in gray. **c**, Box plots showing log2 ratios of the affected (red) and the adjacent normal (gray) regions. Bars within the box plots represent median values. The ends of bars indicate the 25th and 75th percentiles, and the 10th and 90th percentiles are represented by error bars. Student' s *t*-test was performed. NS, not significant. **d**, Copy number and normalized copy number of the 20q11.21 region per diploid genome. Copy numbers were calculated from mean values of log2 ratio in the region. Copy numbers in the homozygous clones were further normalized with the calculated copy number of A1ATD-iPSC line C.



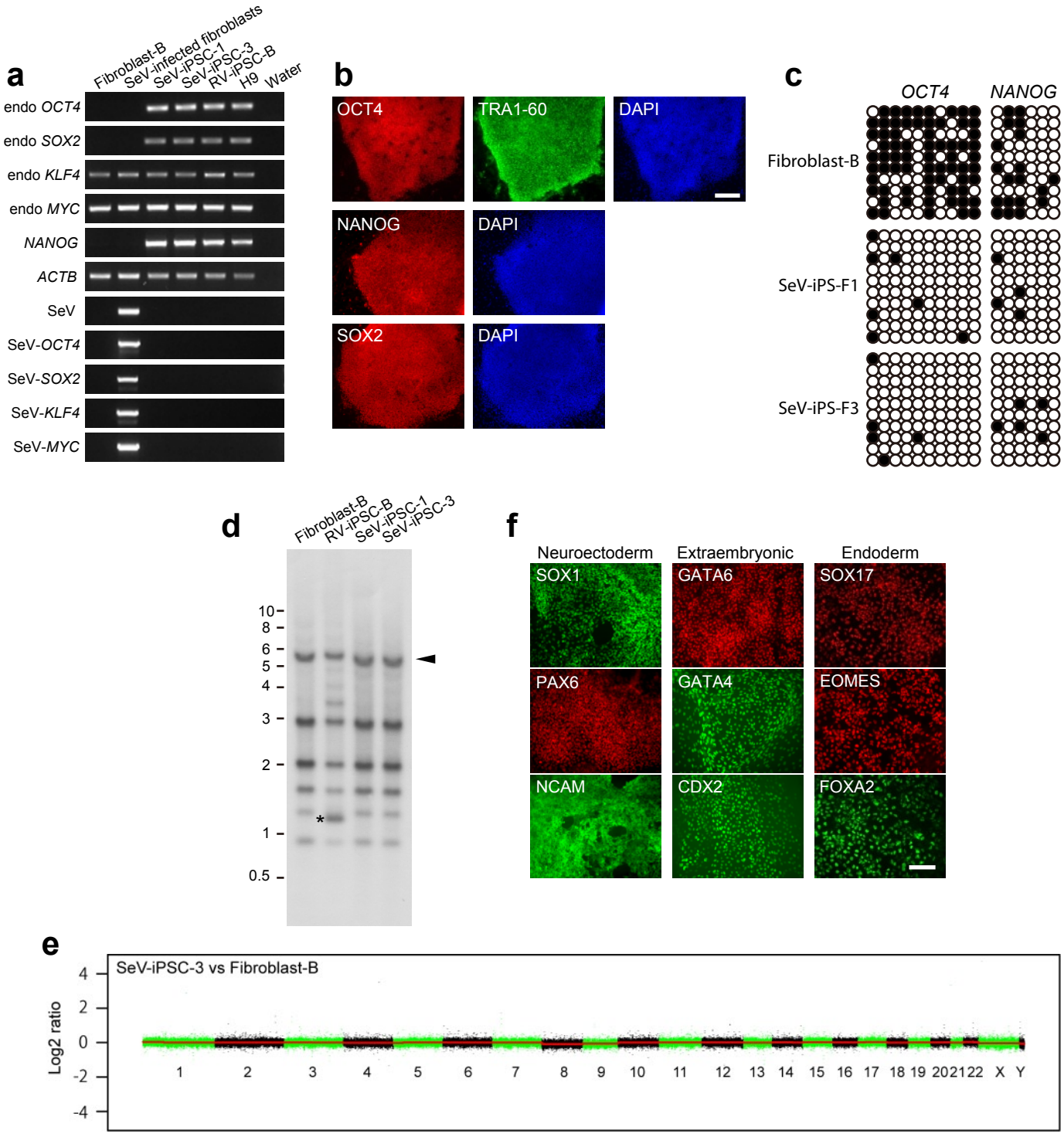
Supplementary Figure 5. SNP analyses.

a, **b**, SNP array data showing entire chromosome 14 (**a**) and 830-kb region around the *A1AT* (*SERPINA1*) gene (**b**). Data for fibroblasts and correctly targeted homozygous clones from each patient are shown. Blue dots represent B allele frequency (left *y* axis) of each SNP and red lines represent smoothed Log *R* ratio (right *y* axis). Note that heterozygosity is maintained in all 6 targeted clones.



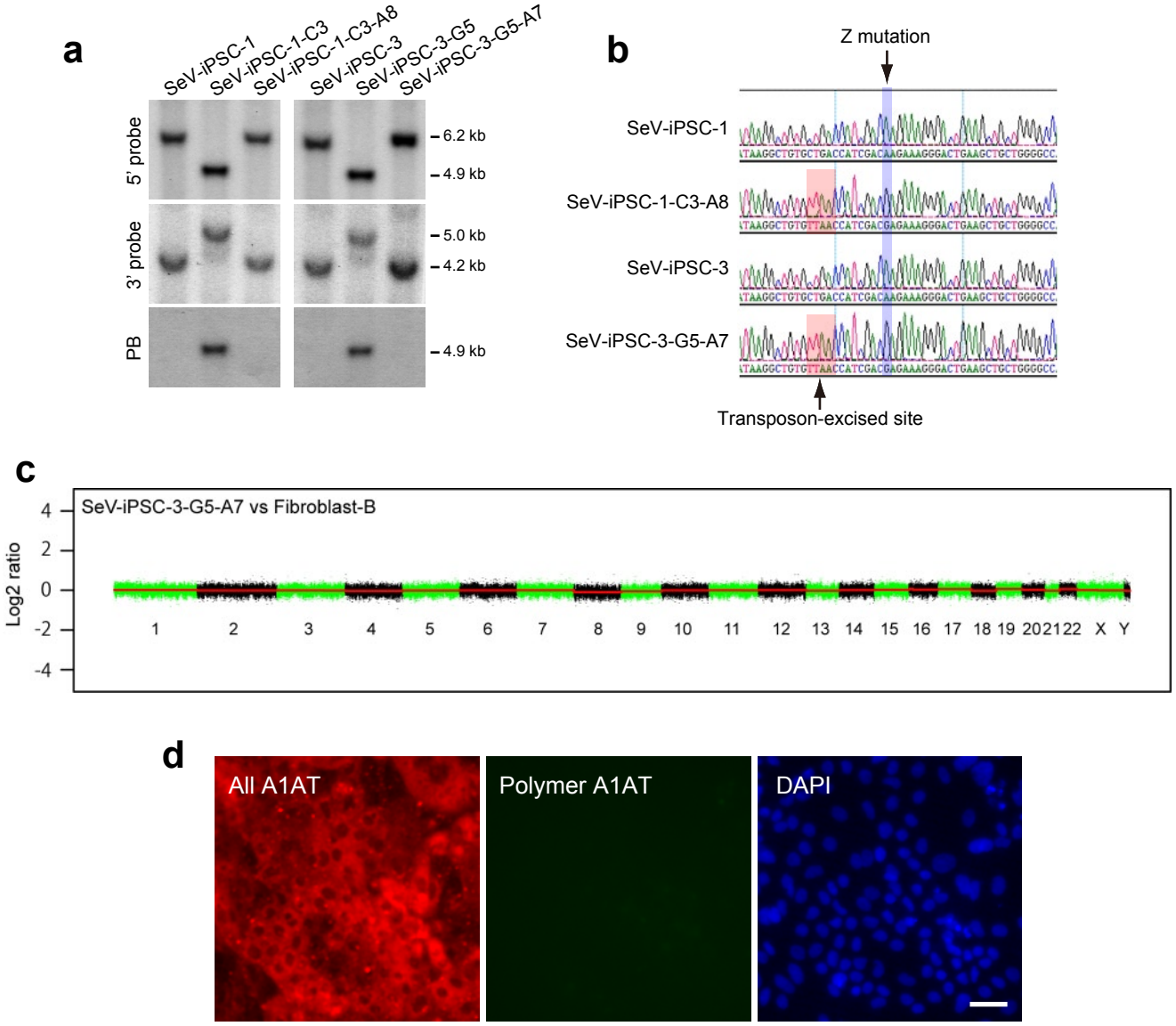
Supplementary Figure 6. Differentiation of corrected hIPSCs into hepatocyte-like cells.

a, Quantitative RT-PCR analyses showing the expression of markers of liver development during c-hIPSCs differentiation. Values are normalized to expression on day 0. **b**, Immunostaining showing the expression of specific proteins marking key stages of hepatocyte development during c-hIPSCs differentiation (Day 3, endoderm; Day 10, hepatic endoderm; Day 18 and 20, hepatic progenitor; Day 25, hepatocyte-like cells). **c**, Flow cytometric analyses showing that 80% of cells are positive for CXCR4 at 3 days of differentiation (endoderm stage) and 80% of cells are positive for Albumin at 25 days of differentiation. **d**-**g**, c-hIPSCs-derived hepatocyte-like cells display functional activity characteristic of primary human hepatocytes including presence of intracellular glycogen storage as shown by periodic acid Schiff staining (**d**), uptake of DiL-LDL (**e**), albumin secretion (**f**) and CytP450 metabolism (**g**). Scale bar, 100 μm



Supplementary Figure 7. Characterization of integration-free A1ATD-hIPSC.

a, **b**, Pluripotency marker expression by RT-PCR (**a**) and immunostaining (**b**) analyses. Note that Sendai virus (SeV) genome were not detected in established iPSC lines. RV, retrovirus. **c**, Bisulfite sequencing of *OCT4* and *NANOG* promoter regions. **d**, Southern blot analysis of the key pluripotency gene, *OCT4*, showing an evidence of integration-free iPSCs. Genomic DNA was digested with PstI. The *OCT4* exon1 was used as a probe. Arrowhead, endogenous *OCT4*; asterisk, provirus-derived *OCT4*. **e**, CGH plot of primary hIPSC line 3, showing no copy number change. **f**, Immunostaining of *in vitro* differentiated cells, showing maintenance of differentiation capacity. Scale bars, 500 μm (b), 200 μm(e)



Supplementary Figure 8. Correction of the Z mutation in the integration-free A1ATD-hIPSC line B. **a**, Validation of genetic modification by Southern blot analysis. See Fig. 2 for the detail. **b**, Sequencing analysis of gene correction. Note that the Z mutation (A) is corrected to wild-type sequence (G), while th *piggyBac* excisions produced TTAA as designed. **c**, CGH analysis showing that the 3-G5-A7 line kept the intact genome compared to the paretal fibroblasts. **d**, Immunostaining of 3-G5-A7-derived hepatocyte-like cells, showing that the corrected cells express normal A1AT porteins. Left, all forms of A1AT; middle, polymeric A1AT; right, DNA staining by DAPI. Scale bar, 50 μm.

Supplementary Table 1. ZFN design

ZFN name	Target sequence [a] (5'-3')	Triplet subsites (5'-3')	Finger designs [b] -1123456	
ZFN-L	gtCGATGGTCAGCAca	CGA	QSAHRIT	F4
		TGG	RSDHLST	F3
		TCA	QSADRTK	F2
		GCA	QSGSLTR	F1
ZFN-R	gaAAGGGActGAAGCTGCTgg	AAG	RLDNRTA	F5
		GGA	QSGHLAR	F4
		GAA	QSGNLAR	F3
		GCT	QSSDLSR	F2
		GCT	QSSDLRR	F1

[a] The core DNA target sequences (capital letters) and 2 flanking bases on each side (small letters) are shown. [b] The amino acid residues at positions '-1' to '+6' of the recognition alpha helix[31,32] of each of the zinc finger DNA-binding domain for each DNA triplet target are shown.

Supplementary Table 2. aCGH analyses of primary A1ATD-hIPSCs and their derivatives

Cell line	*A1AT* genotype	Reference Genome	Abnormality	Chr.	Band	Size	Mean log2 ratio	No. of affected genes
a. Primary A1ATD-hIPSCs								
A	*Z/Z*	A-Fibroblast	-					
B	*Z/Z*	B-Fibroblast	Gain	1	p36.11	300 kb	0.60	8
			Loss	8	q22.2	20 kb	-0.93	1
C	*Z/Z*	C-Fibroblast	Loss	2	p22.3	207 kb	-0.96	1
			Loss	17	q23.1	142 kb	-0.76	3
			Gain *	20	q11.21	1.34 Mb	0.21	30
b. Homozygously targeted clones								
A-A9	*PB/PB*	A	-					
A-G6	*PB/PB*	A	Loss	14	q23.13	23 kb	-0.87	0
B-16	*PB/PB*	B	-					
C-B6	*PB/PB*	C	Gain *	20	q11.21	1.34 Mb	0.71	30
C-C11	*PB/PB*	C	-					
C-G4	*PB/PB*	C	-					
c. Clones with bi-allelic *piggyBac* excision								
B-16-C2	*Rev/Rev*	B	-					
B-16-C3	*Rev/Rev*	B	-					
B-16-C6	*Rev/Rev*	B	-					
B-16-D2	*Rev/Rev*	B	-					
B-16-E4	*Rev/Rev*	B	Gain	22	q11.23	47 kb	0.89	1
B-16-E5	*Rev/Rev*	B	Gain	15	q26.2	4.7 kb	1.15	1
B-16-F4	*Rev/Rev*	B	-					
B-16-F10	*Rev/Rev*	B	Gain	17	q23.3-q24.2	4.7 Mb	0.62	56
B-16-G8	*Rev/Rev*	B	-					
B-16-G10	*Rev/Rev*	B	Gain	4	q31.3	262 kb	0.57	1
B-16-H4	*Rev/Rev*	B	-					
C-G4-B3	*Rev/Rev*	C	-					
C-G4-C2	*Rev/Rev*	C	-					
C-G4-C10	*Rev/Rev*	C	-					
C-G4-D7	*Rev/Rev*	C	-					
C-G4-E9	*Rev/Rev*	C	-					
d. *in vitro* differentiated hepatocytes								
B-16-C2-hep.	*Rev/Rev*	B-16-C2	-					

* See Supplementary Fig. 4 and Supplementary Analysis for more detail.

Supplementary Table 3. Validated mutations in protein coding regions indentified by exome sequencing

Chr	Position	Wt Allele	Mut Allele	Gene_ID	Transcript_ID	Mutation type	cDNA annotation	Protein annotation	Detected in				SIFT functional prediction	Expression in hepatocytes ***
									Fibro-B	iPSC-B	B-16	B-16-C2		
Chr1	35223302	G	A	GJB5	ENST00000338513_r62	MISSENSE	c.371G>A	p.R124Q					TOLERATED	-
Chr1	173499158	C	A	SLC9A11	ENST00000367714_r62	SILENT	c.2199G>T	p.V733V					N/A	-
Chr1	233314872	C	A	PCNXL2	ENST00000258229_r62	MISSENSE	c.3116G>T	p.G1039V					DAMAGING	-
Chr3	42705381	C	T	ZBTB47	ENST00000457842_r62	SILENT	c.702C>T	p.G234G					N/A	+
Chr3	100962755	C	T	IMPG2	ENST00000193391_r62	MISSENSE	c.2420G>A	p.R807K					TOLERATED	-
Chr3	189586471	G	A	TP63	ENST00000264731_r62	SILENT	c.1095G>A	p.S365S					N/A	-
Chr3	189586472	G	A	TP63	ENST00000264731_r62	MISSENSE	c.1096G>A	p.D366N					TOLERATED	
Chr4	2831347	G	A	SH3BP2	ENST00000452765_r62	SILENT	c.714G>A	p.P238P					N/A	+
Chr4	57182264	C	A	KIAA1211	ENST00000264229_r62	MISSENSE	c.2596C>A	p.Q866K					TOLERATED	-
Chr4	122254096	G	T	QRFPR	ENST00000394427_r62	MISSENSE	c.677C>A	p.P226H					DAMAGING	-
Chr5	14692993	C	T	FAM105B	ENST00000284274_r62	MISSENSE	c.895C>T	p.R299C					TOLERATED	+
Chr5	33936907	T	C	RXFP3	ENST00000330120_r62	MISSENSE	c.62T>C	p.L21P					DAMAGING*	-
Chr6	166883328	C	A	RPS6KA2	ENST00000265678_r62	SPLICE	Exon11+1G>T	N/A					N/A	+
Chr7	5567522	T	C	ACTB	ENST00000331789_r62	MISSENSE	c.985A>G	p.I329V					DAMAGING*	+
Chr9	124083640	A	C	GSN	ENST00000373823_r62	MISSENSE	c.1286A>C	p.Y429S					DAMAGING	+
Chr9	125424828	C	T	OR1L1	ENST00000373686_r62	SILENT	c.984C>T	p.T328T					N/A	-
Chr11	4143385	G	A	RRM1	ENST00000300738_r62	MISSENSE	c.1053G>A	p.M351I					DAMAGING	+
Chr11	28119199	C	T	KIF18A	ENST00000263181_r62	MISSENSE	c.296G>A	p.R99H					DAMAGING	-
Chr11	76909568	C	A	MYO7A	ENST00000409709_r62	SILENT	c.4470C>A	p.I1490I					N/A	+
Chr12	129180382	C	T	TMEM132C	ENST00000315208_r62	MISSENSE	c.511C>T	p.R171C					DAMAGING	-
Chr13	99361826	G	A	SLC15A1	ENST00000376503_r62	MISSENSE	c.1067C>T	p.T356I					DAMAGING	+
Chr14	69521806	A	G	DCAF5	ENST00000341516_r62	MISSENSE	c.1597T>C	p.S533P					DAMAGING	+
Chr14	94844957	C	T	SERPINA1 **	ENST00000440909_r62	SILENT	c.1086G>A	p.L362L					N/A	+
Chr14	94844959	G	A	SERPINA1 **	ENST00000440909_r62	SILENT	c.1084C>T	p.L362L					N/A	
Chr14	106378125	G	T	IGHD5-5	ENST00000390588_r62	MISSENSE	c.11C>A	p.A4D					Not scored	-
Chr15	68497606	C	T	CALML4	ENST00000448060_r62	MISSENSE	c.109G>A	p.G37S					TOLERATED	+
Chr15	86225404	A	C	AKAP13	ENST00000361243_r62	MISSENSE	c.5129A>C	p.H1710P					DAMAGING	+
Chr17	79428053	C	G	BAHCC1	ENST00000307745_r62	MISSENSE	c.6364C>G	p.L2122V					TOLERATED	-
Chr18	52899761	G	-	TCF4	ENST00000356073_r62	FRAMESHIFT	c.1628 C>-	p.S543fs*1					N/A	+
ChrX	12938569	C	A	TLR8	ENST00000218032_r62	MISSENSE	c.1410C>A	p.F470L					TOLERATED	-
ChrX	107823778	C	A	COL4A5	ENST00000328300_r62	SILENT	c.796C>A	p.R266R					N/A	-

N/A, not applicable

* Low confidence

** These two silent mutations were intensionally introduced though genetic corretion.

*** Expression in hepatocyte was analyzed using control samples in GSE11942. Genes with 'present' call in more than 50% of samples are considered as expressed.

Not detected

Detected

Supplementary Table 4. aCGH analyses of integration-free iPSCs and their derivatives.

Cell line	*A1AT* genotype	Reference Genome	Abnormality	Chr.	Band	Size	Mean log2 ratio	No. of affected genes
a. Primary factor-free A1ATD-hIPSCs								
SeV-iPSC1	*Z/Z*	Fibroblast-B	Gain	6	q22.1	271kb	0.52	2
SeV-iPSC2	*Z/Z*	Fibroblast-B	Loss	2	q23.3	18kb	-0.93	1
SeV-iPSC3	*Z/Z*	Fibroblast-B	-					
SeV-iPSC4	*Z/Z*	Fibroblast-B	Gain	10	q21.1	1.1Mb	0.52	2
			Gain	16	p13.12	76kb	0.61	0
			Gain	22	q13.33	86kb	0.63	3
SeV-iPSC5	*Z/Z*	Fibroblast-B	Gain	18	p11.31-p11.23	951kb	0.6	5
SeV-iPSC6	*Z/Z*	Fibroblast-B	Loss	22	q13.1	10 kb	-0.81	1
b. Homozygously targeted clones								
iPSC1-C3*	*PB/PB*	Fibroblast-B	-					
iPSC3-G5	*PB/PB*	Fibroblast-B	-					
c. Clones with bi-allelic *piggyBac* excision								
iPSC1-C3-A8*	*Rev/Rev*	Fibroblast-B	-					
iPSC1-C3-A9*	*Rev/Rev*	Fibroblast-B	-					
iPSC3-G5-A7	*Rev/Rev*	Fibroblast-B	-					
iPSC3-G5-C8	*Rev/Rev*	Fibroblast-B	Loss	20	p12.1	421kb	-1.00	1

* A deletion on Chr.6 found in the parental line is not shown.

Supplementary Table 5. Primer sequences

Construction of pPB-R1R2-NP	attR1-F	CTAGCTAGCACAAGTTTGTACAAAAAAGCTGAAC
	attR1-R	CCCAAGCTTGAATTCGGATCCCATAGTGACTGGATATGTTGTGTT
	attR2-F	CCGCTCGAGTCTAGACATAGTGACTGGATATGTTGTGTT
	attR2-R	GACTAGTACCACTTTGTACAAGAAAGCTGAA
	EM7neo-F	CCCAAGCTTGTTGACAATTAATCATCGGCATAG
	EM7neo-R	CCGCTCGAGTCAGAAGAACTCGTCAAGAAGGCG
Construction of the Tyr-mini targeting vector	Left arm-F	GGCGCGCCATGGAACAGTGAAGTTCTCATCCCCAG
	Left arm-R	TTTAACGTACGTCACAATATGATTATCTTTCTAGGGTTAATGATATCAACATCTACGACCTCTTTGTATG
	Right arm-F	AAGAATGCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAAACATAGGTGTTGATCCATTGTTCATTTGGCCATA
	Right arm-R	CCTTAATTAAGCGGAAACTGTAAGTTTGGATTTG
Construction of the Tyr-retrieving vector	Left arm-F	CCGCTCGAGTGCCCATGCCTAAACACAGATGTA
	Left arm-R	CCCAAGCTTTTTTGGATTCGGATTCTGAGGCT
	Right arm-F	CCCAAGCTTGTACAAGATTTTCATATGTAAGC
	Right arm-R	AGGCGCGCCAGGAATTCATGCCCCAGTTGACA
Homologous recombination detection at Tyr	F	GTTGGCGCCTACCGGTGGATGTGGAATGTG
	R	GAAATCTCTTCAAGCTGGGAAAACCTAAAG
Tyr genotyping	P1	TCCTCGAGCCTGTGCCTCCTCTAAG
	P2	CCCAAGTACTCATCTGTGCAAATGTC
	P3	GCGACGGATTCGCGCTATTTAGAAAG
Tyr 5' probe	F	ACTCATTATTCCAGGATACTTGAGTG
	R	ATTTACCACATGCCCAAACTAATAAC
A1AT 2-kb fragment	F	GTGGTGGGTCCCAGAAGAACAAGAGGAATGC
	R	CATAGCTGAGGAGTCCTTGCAATGGCCTTCC
Sequencing primers of the A1AT 2-kb region	Seq1	AGGGGCCGAGGGAAACAAATGAAGA
	Seq2	GCTAAAGATGACACTTATTTTGGAA
	Seq3	ATGTGACAGGGAGGGAGAGGATGTG
	Seq4	GAAAAGTGGTGAATCCCACCCAAA
	Seq5	GCCCATCTGTTTCTGGAGGGCTCCA
	Seq6	TCACCTAACCAGACTCGGGCCCTGC
	Seq7	TACCAGGGTGCAACAAGGTCGTCAG
Construction of the A1AT-donor template vector	Left arm-F	GGCGCGCCTCTGCACGACAGGTCTGCCAGCTTAC
	Left arm-R	AAGAATGCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAACACAGCCTTATGCACGGCCTGGAGGGGAG
	Right arm-F	TTAACGTACGTCACAATATGATTATCTTTCTAGGGTTAACCATCGACGAGAAAGGGACTGAAGCTGCTGG
	Right arm-R	CCTTAATTAATGGTCTTCTGGGGCCTGCTGGGC
Genotyping after ZFN-mediated targeting	U4	TGGAGTGACGATGCTCTTCCCTGTTC
	L1	GGTCAATGGGTGATGTGCTTCCTCTC
	PB5P2	CGTCAATTTTACGCATGATTATCTTTAAC
Detection of backbone integration	F	ACACAGGAAACAGCTATGACCATGATTACG
	R	CGTCAATTTTACGCATGATTATCTTTAAC
Genotyping after transposon excision	U4	TGGAGTGACGATGCTCTTCCCTGTTC
	L2	GCAGTTATTTTTGGGTGGGATTCAC
	PB3-P2	GCGACGGATTCGCGCTATTTAGAAAG
Detection of transposon re-integration	F	CTGCTGCAACTTACCTCCGGGATG
	R	CCAATCCTCCCCCTTGCTGTCCTG
A1AT 5' probe	F	GATTCCCCAACCTGAGGGTGACCAAG
	R	AAGTGACAGAGAAACGCAAGCCTTC
A1AT 3' probe	F	CGAGGGTCACGCTAAACTTCTGCAG
	R	CTCTGGGAGGTTATGCTAGTTTGAA
Surveyor Nuclease assay	F	AGGAGCAAGGCCTATGTGA
	R	GAGGAGCGAGAGGCAGTTAT

Supplementary Analysis

Analyses on 20q11.21 amplification

In the initial analysis of CGH data, we found amplification of 20q11.21 in one homozygous clone, C-B6 (Supplementary Fig. 4a). Through analysis of SNP array data, however, this amplification was found in all lines derived from A1ATD-hIPSC line C (Supplementary Fig. 4b). This region is found to be frequently amplified following prolonged hESC culture[1,2]. We individually analyzed CGH data of the 20q11.21 region in 3 primary A1ATD-hIPSC lines and 3 homozygous clones derived from line C. We first compared log2 ratios in the affected region (position 29,297,270 - 30,638,579; 167 probes) and those in the adjacent normal copy number region (position 30,642,607 - 32,056,016) that contains the same number of probes (Supplementary Fig. 4a). Consistent with the SNP array data, all three homozygous clones showed significantly higher ratios in the affected region than in the adjacent region (Supplementary Fig. 4c). However, the ratios in C-C11 and C-G4 were lower than the cut-off used for the initial CGH data analysis. C-B6 showed a much higher ratio than C-C11 and C-G4, suggesting that C-B6 carried more copies of the 20q11.21 region than the others. Surprisingly, the primary hIPSC line C also showed a slightly higher ratio in the affected region than in the adjacent region whereas line A and B, both carrying normal chromosome 20 based on the SNP analysis, had same ratios. This result strongly suggests that line C is heterogeneous with a minor population of cells with amplified 20q11.21. The mean log2 ratio of this region in line C is 0.208, which indicates 2.3 copies per diploid genome (Supplementary Fig. 4d). On the assumption that line C consists of 2 cell types (2-copy normal and 3-copy abnormal cells), 30 % of the population carries 3 copies of 20q11.21. Since we used genomic DNA from line C as a reference for subsequent CGH analyses of the line C-derived cell lines, the log2 ratio of the 20q11.21 region in C-C11 and C-G4 was lower than a ratio representing single-copy gain (0.58) and thus this region was not called as a “Gain”. We recalculated the copy number for 3 homozygous clones using the corrected copy number for the line C. Normalized copy numbers of 20q11.21 were 3 copies for C-C11 and C-G4, and 3.8 for C-B6 (Supplementary Fig. 4d), which is consistent with the SNP array data. Thus, in C-B6, the 20q11.21 region was further amplified. The ZFN-stimulated gene targeting in line C was conducted several passages after CGH analysis. The abnormal cells might become dominant at the time of gene targeting. As a result all derivative lines carried 20q11.21 amplification.

References

1 Lefort, N. *et al.* Human embryonic stem cells reveal recurrent genomic instability at 20q11.21. *Nat Biotechnol* **26**, 1364-1366 (2008).

2 Spits, C. *et al.* Recurrent chromosomal abnormalities in human embryonic stem cells. *Nat Biotechnol* **26**, 1361-1363 (2008).

Analysis on one point mutation introduced during ZFN-stimulated targeting

As ZFN-induced mutations are primarily InDels, the one observed single base substitution is unlikely to be the result of off-target ZFN cleavage. To support this conclusion we scanned the human genome (hg19) for potential ZFN cleavage sites and calculated the distance between maximum-likelihood off-target sites and the mutation. Both heterodimers and homodimers of either site with 5 or 6bp between individual ZFN sites and up to 6 mismatches from the intended binding sites were allowed; this returned 495 potential cleavage sites. The smallest distance between any of these 495 potential cleavage sites and any of the mutation was 110,407 bp. Thus, it is extremely unlikely that ZFN cleavage at any of these putative off-target sites caused the observed mutation.

Supplementary Note. Amino acid sequences of ZFNs against the *Z* mutation in *A1AT*.
Amino acid sequences of left (L) and right (R) ZFNs are shown. Zinc finger protein regions are underlined.

ZFN-L

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAMAERPFQCRICMRNFSQSGSLTRHIRTHTG
EKPFACDICGRKFAQSADRTKHTKIHTGSQKPFQCRICMRNFSRSDHLSTHIRTHTGEKPFACDICGRKF
AQSAHRITHTKIHLRGSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVY
GYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWK
VYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGE
INFRS

ZFN-R

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAMAERPFQCRICMRKFAQSSDLRRHTKIHTG
EKPFQCRICMRNFSQSSDLSRHIRTHTGEKPFACDICGRKFAQSGNLARHTKIHTPNPHRRTDPSHKPFQ
CRICMRNFSQSGHLARHIRTHTGEKPFACDICGRKFARLDNRTAHTKIHLRGSQLVKSELEEKKSELRHK
LKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKA
YSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITN
CNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINFRS