Supplementary Table 1. Variant numbers in different QC stages. S1: SNP-set after genotyping rate, missing rate in samples and removal of monomorphic sites. S2: removal of non-loss-of-functional related SNPs from S1 dataset.

Supplementary Table 2. All nonsynomymous SNVs in *FGF6* exon regions and missense mutation prediction with different algorithms.

Supplementary Table 3. Real-time PCR Primers for FGF-6 network validation.

Supplementary Figure Legends

Fig. S1. Two-Site Power Calculations. Power calculations for a two-site disease model comparing the Armitage trend test of disease association at each site to a log-likelihood ratio test explicitly evaluating recessive diplotype effects. Baseline haplotype frequencies, case and control diploid sample sizes, and relative risk of disease-predisposing diplotypes parameters are shown. The initial haplotype frequencies (A1B1, A1B2, A2B1, A2B2) are presented. Different combinations of haplotypes are generated by generating recombination between the two sites and the results are presented in a collapsed manner through a single linkage disequilibrium metric. Hardy-Weinberg equilibrium of haplotypes/diplotypes in the general population is assumed. R is the relative risk of disease for recessive diplotypes compared to the remaining diplotypes. ncs and nct are the number of cases and controls, respectively. The type I error rate, adjusted for an exome-wide scan, was set to 2.5×10^{-6} for all calculations.

Fig. S2. Minor Allele Frequency Distribution to PMRP dataset. Displayed is the histogram of the minor allele frequency (MAF) at each variant within the 10,000 PMRP subjects following removal of variants from the QC procedures.

Fig. S3. Quantile-Quantile Plot. Q-Q plot for the exome-wide, gene-based recessive diplotype scanning in hemochromatosis is shown. Numerous genes had no recessive diplotypes with putative functional alleles and therefore yielded P-values of 1. The two data points exceeding the confidence interval represented *HFE* and *FGF6*.

Fig. S4. Comparative genomic analysis and protein-protein interaction (PPI). The comparative genomic analyses revealed that *FGF6* evolved synchronously with other iron metabolism genes. **(A)** Main iron metabolism genes were collected and alignment was conducted to make the comparative genomic analysis together with *FGF6*. The earliest gene appearance over time was inferred by comparing species and corresponding evolution and appearance time was labelled. **(B)** Protein-protein interaction network was estimated by String (version 10.0)[51] using the highest confidence setting (confidence score>0.9).

Fig. S5. Total iron content in HFF-1 cells with increasing FGF-6 protein concentration. * P＜.05; ** P＜ 0.01. Results are the mean±SD of 3 observations in each experiment.

Fig. S6. FGF6 mutation Plasmid Structures in the study. M1 (GAG->TAG) E172X, M2 (GAC->GTC) D174V and M3 (CGG -> CAG) R188Q

Fig. S7. Relative transcription of iron metabolism genes with wildtype FGF6 and three mutations. mRNA levels of FGF6 and four iron metabolism genes are measured relative to GAPDH following transfection of vector, wildtype FGF6 (WT), E172X FGF6 (M1), D174V FGF6 (M2), and R188Q FGF6 (M3). Measurements were taken in three cell lines: 786-O, A498, and HCT-8.

Fig. S8. FGF6 loss-of-function nonsynonymous variants cause hepcidin downregulation and iron deposition in HFF-1. (A) Iron metabolism gene expression changes after the transfection by *FGF6* mRNA into HFF-1 with wildtype and the identified variants R188Q, D174V and E172X. **(B)** Total iron contents changes after the transfection by *FGF6* mRNA into HFF-1 with wildtype and the identified variants R188Q, D174V and E172X. **(C)** Ferritin protein level changes after the transfection by *FGF6* mRNA into HFF-1 with wildtype and the identified

variants R188Q, D174V and E172X. **(D)** The densitometry data of Western blot for Ferritin protein were shown in the column chart. *$P<.05$; **$P<0.01$. Results are the mean±SD of three observations in one experiment.

Fig. S9. Perls' stain reveals that FGF6 loss-of-function nonsynonymous variants cause iron deposition. Perls' stain in HepG2 **(A)**, HCT-116 **(B)**, HCT-8 **(C)**, 786-O **(D)** and HFF-1 **(E)** in the presence of FAC differs among transfection by *FGF6* mRNA with wildtype and the identified variants R188Q, D174V and E172X.

Fig. S10. Perls' stain in SSc and liver cancer. (A) Perls' stain was applied to evaluate the iron deposition in SSc skin tissues. Perls' stain was visualized by Nikon microscopy. The ratio of iron-positive stain areas to the total area was used to evaluate the iron deposition levels by Image J software. Arrows indicated positive stain area. **(B)** Perls' stain in liver cancer tissues. Perls' stain was visualized by Nikon microscopy. The ratio of iron-positive stain areas to the total area was used to evaluate the iron deposition levels by Image J software.

Fig. S11. FGF6 protein levels were different among normal, cancer and metastatic cells. (A) IHC of FGF-6 in normal hepatocytes and metastatic cells. The blue circle indicated normal liver tissue and the arrows indicated metastatic cells. (B) IHC of FGF-6 in non-metastatic liver cancer cells.

Supplemental Reference

SR1. Szklarczyk, D. et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 43, D447-52 (2015).

Supplementary Table 1. Variants numbers in different QC stages

		S1	S2	
MAF Internal	No#	Freq%		
0	126400	0.221892582		
0-0.001	99512	0.174691255	84360	
0.001–0.005	24629	0.043235699	18428	
0.005–0.01	7717	0.013547034	5023	
0.01–0.05	22547	0.039580792	8102	
>0.05	259296	0.455188758	13643	
	413701		129556	P=0.22

S1: SNP-set after genotyping rate, missing rate in samples and monopolymorphism removing.
S2: remove non-loss-of-functional related SNPs from S1 dataset.

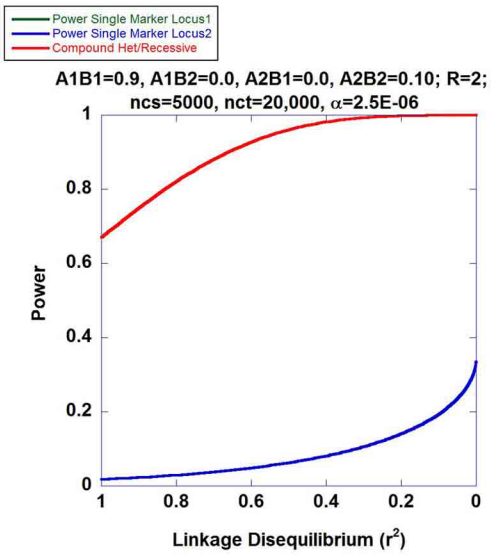
S. Table 2. All nonsynomymous SNV in FGF6 exon regions and missense mutation prediction with different algorithms.

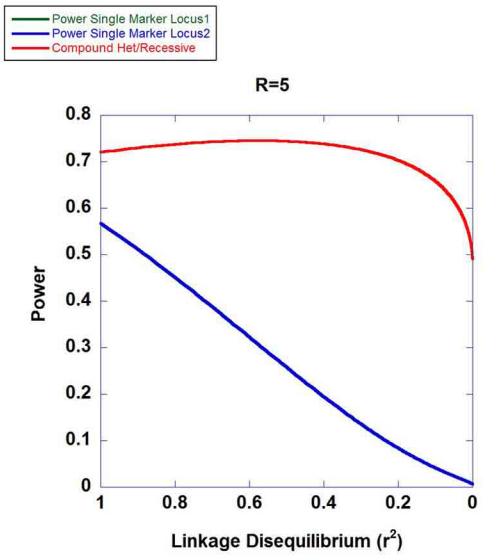
Chr	Start	End	Alt	Func.refGene	Gene.refGene	GeneDetail.refGene	AAChange.refGene	phastConsElements46way	Polyphen2_HDIV	Polyphen2_HVAR	LRT	MutationTaster	MutationAssessor	FATHMM	PROVEAN	VEST3	MetaLR	M-CAP	CADD
12	4543445	C	T	exonic	FGF6	nonsynonymous SNV	FGF6:NM_020996:exon3:c.G563A:p.R188Q	D	D	D	D	D	L	T	N	T	T	D	D
12	4543487	T	A	exonic	FGF6	nonsynonymous SNV	FGF6:NM_020996:exon3:c.A521T:p.D174V	T	B	B	N	P	N	T	N	T	T	.	N
12	4543494	C	A	exonic	FGF6	stopgain	FGF6:NM_020996:exon3:c.G514T:p.E172X	.	.	.	D	D	.	.	.	.	.	.	D
12	4543523	G	A	exonic	FGF6	nonsynonymous SNV	FGF6:NM_020996:exon3:c.C485T:p.T162I	T	B	B	N	D	N	D	N	T	T	T	D
12	4553300	G	A	exonic	FGF6	nonsynonymous SNV	FGF6:NM_020996:exon2:c.C449T:p.T150M	D	D	D	D	D	M	D	D	D	D	D	D
12	4553304	C	T	exonic	FGF6	nonsynonymous SNV	FGF6:NM_020996:exon2:c.G445A:p.A149T	T	B	B	N	N	L	D	N	T	T	D	D
12	4553361	G	C	exonic	FGF6	nonsynonymous SNV	FGF6:NM_020996:exon2:c.C388G:p.L130V	D	D	P	D	D	M	D	N	D	D	D	D
12	4553391	T	C	exonic	FGF6	nonsynonymous SNV	FGF6:NM_020996:exon2:c.A358G:p.I120V	D	P	P	D	D	L	D	N	D	D	D	D
12	4554550	C	T	exonic	FGF6	nonsynonymous SNV	FGF6:NM_020996:exon1:c.G187A:p.A63T	T	B	B	N	N	L	T	N	T	T	.	N
12	4554630	A	G	exonic	FGF6	nonsynonymous SNV	FGF6:NM_020996:exon1:c.T107C:p.V36A	T	B	B	N	P	L	T	N	T	T	.	D
12	4554651	C	A	exonic	FGF6	nonsynonymous SNV	FGF6:NM_020996:exon1:c.G86T:p.G29V	D	P	B	D	D	L	T	N	T	T	T	D

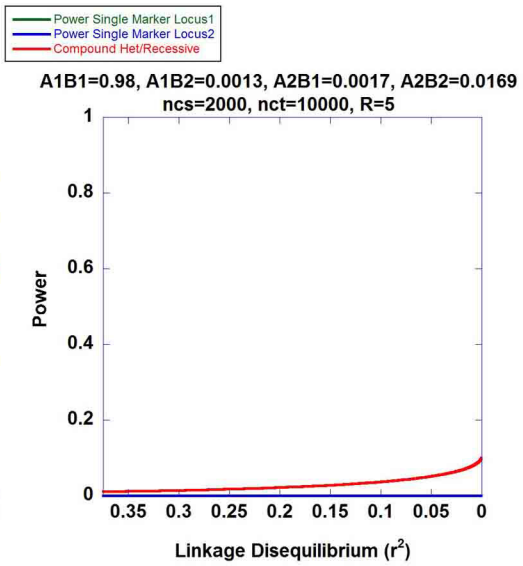
Supplementary Table 3. Real-time PCR Primers for FGF6 network validation

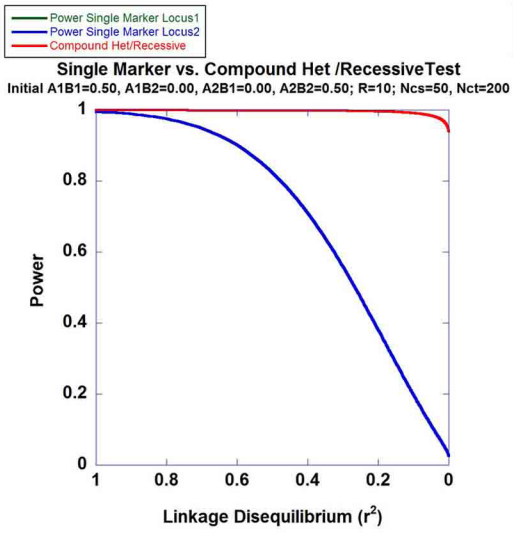
No.	Gene symbol	Accession number	Forward primer	Reverse primer	Forward TM	Reverse TM	Amplicon size
1	HAMP	NM_021175.3	CTGACCAGTGGCTCTGTTTTCC	AAGTGGGTGTCTCGCCTCCTTC	61.39	64.21	128
5	HDAC2	NM_001527.3	AGCCACTGCCGAAGAAATGA	AACTTCACAGCTCCAGCAACT	59.68	59.86	198
24	GAPDH	NM_001289745.2	GAGTCAACGGATTTGGTCGT	CATGGGTGGAATCATATTGGA	58.21	55.39	141
25	FGF6	NM_020996.2	GTGCCCTCTTCGTTGCC	TACCCGTCCGTATTTGCT	58.3	55.58	164
26	HEPH	NM_001130860.3	CAGGTGGTCTTCTACAACCGTG	TGGCAACCAAGCCAGGGTAAGA	60.87	64.17	124
27	TFRC	NM_001313966.1	ATCGGTTGGTGCCACTGAATGG	ACAACAGTGGGCTGGCAGAAAC	63.42	63.8	131

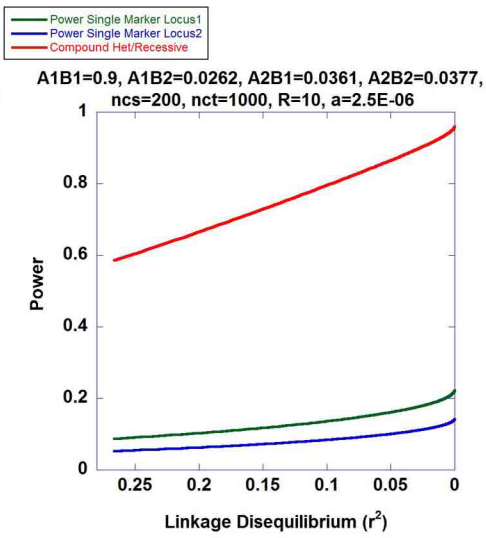
Figure S1

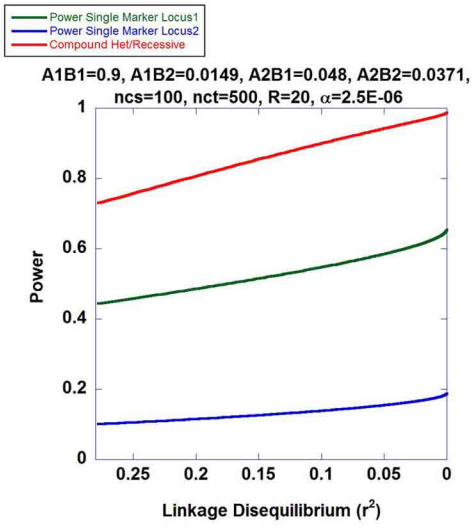

Power Single Marker Locus1
Power Single Marker Locus2
Compound Het/Recessive
A1B1=0.9, A1B2=0.0, A2B1=0.0, A2B2=0.10; R=2;
ncs=5000, nct=20,000, α=2.5E-06
Power
Linkage Disequilibrium (r2)

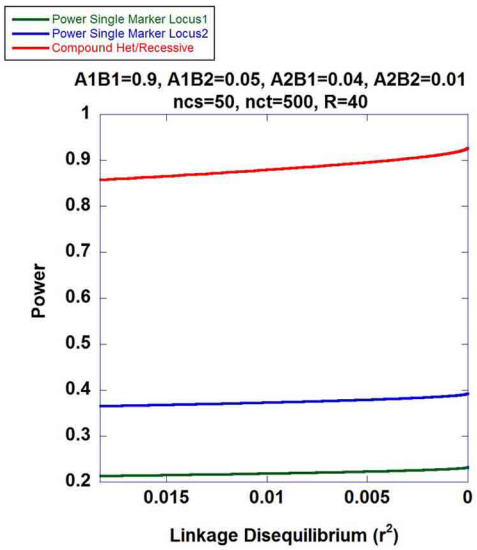

Power Single Marker Locus1
Power Single Marker Locus2
Compound Het/Recessive
R=5
Power
Linkage Disequilibrium (r2)

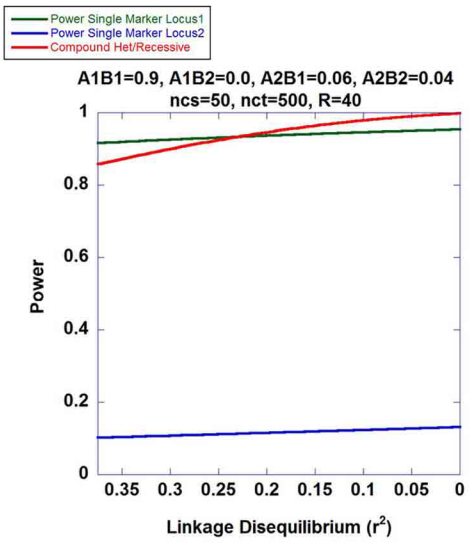

Power Single Marker Locus1
Power Single Marker Locus2
Compound Het/Recessive
A1B1=0.98, A1B2=0.0013, A2B1=0.0017, A2B2=0.0169
ncs=2000, nct=10000, R=5
Power
Linkage Disequilibrium (r2)


Power Single Marker Locus1
Power Single Marker Locus2
Compound Het/Recessive
Single Marker vs. Compound Het /RecessiveTest
Initial A1B1=0.50, A1B2=0.00, A2B1=0.00, A2B2=0.50; R=10; Ncs=50, Nct=200
Power
Linkage Disequilibrium (r2)


Power Single Marker Locus1
Power Single Marker Locus2
Compound Het/Recessive
A1B1=0.9, A1B2=0.0262, A2B1=0.0361, A2B2=0.0377,
ncs=200, nct=1000, R=10, a=2.5E-06
Power
Linkage Disequilibrium (r2)


Power Single Marker Locus1
Power Single Marker Locus2
Compound Het/Recessive
A1B1=0.9, A1B2=0.0149, A2B1=0.048, A2B2=0.0371,
ncs=100, nct=500, R=20, α=2.5E-06
Power
Linkage Disequilibrium (r2)


Power Single Marker Locus1
Power Single Marker Locus2
Compound Het/Recessive
A1B1=0.9, A1B2=0.05, A2B1=0.04, A2B2=0.01
ncs=50, nct=500, R=40
Power
Linkage Disequilibrium (r2)


Power Single Marker Locus1
Power Single Marker Locus2
Compound Het/Recessive
A1B1=0.9, A1B2=0.0, A2B1=0.06, A2B2=0.04
ncs=50, nct=500, R=40
Power
Linkage Disequilibrium (r2)

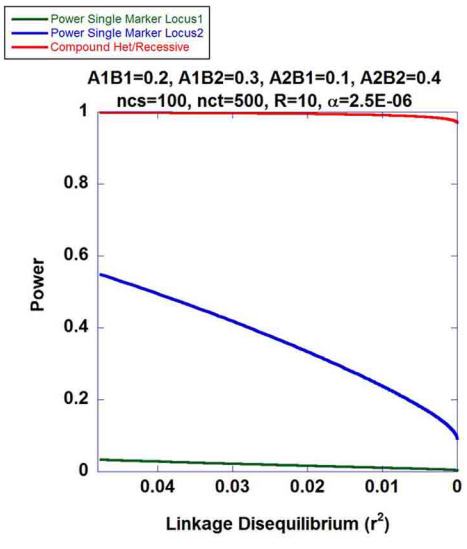

Power Single Marker Locus1
Power Single Marker Locus2
Compound Het/Recessive
A1B1=0.2, A1B2=0.3, A2B1=0.1, A2B2=0.4
ncs=100, nct=500, R=10, α=2.5E-06
Power
Linkage Disequilibrium (r2)

Figure S2

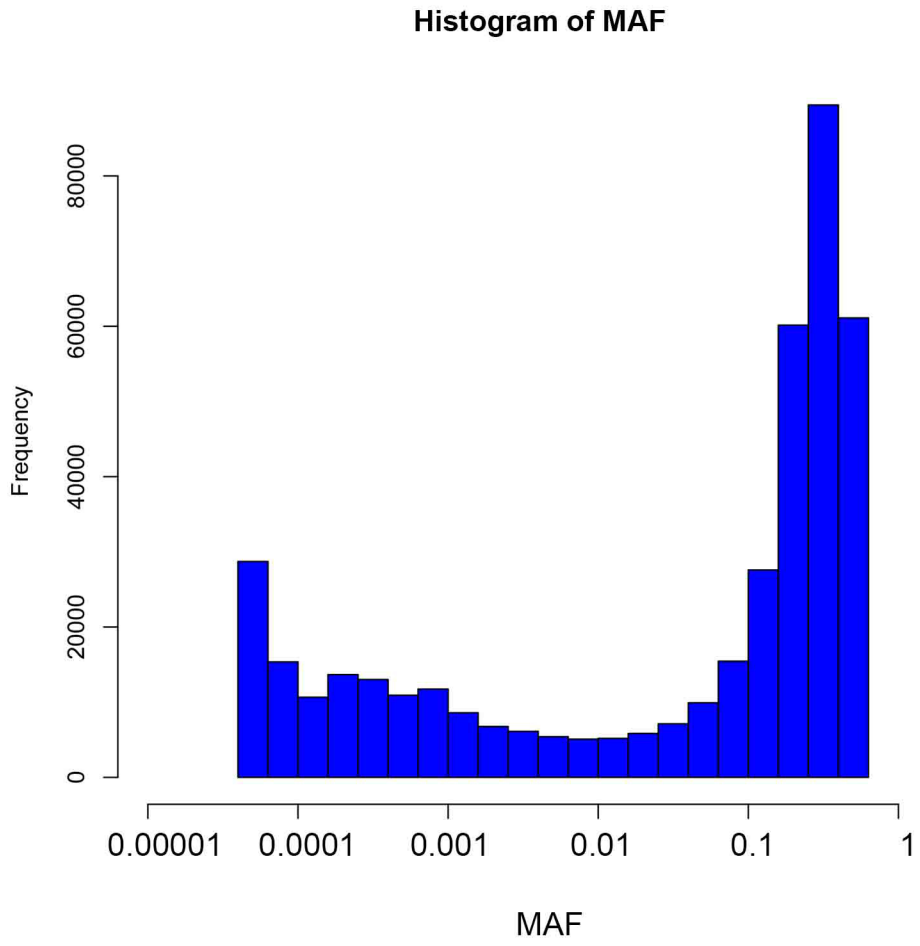

Histogram of MAF
Frequency
0
20000
40000
60000
80000
0.00001
0.0001
0.001
0.01
0.1
1
MAF

Figure S3

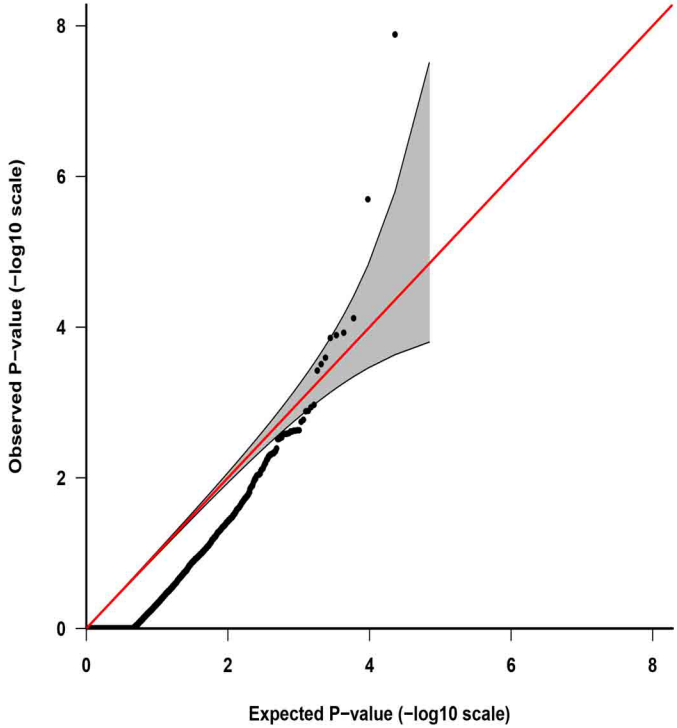

Observed P−value (−log10 scale)
Expected P−value (−log10 scale)
0
2
4
6
8

Figure S4

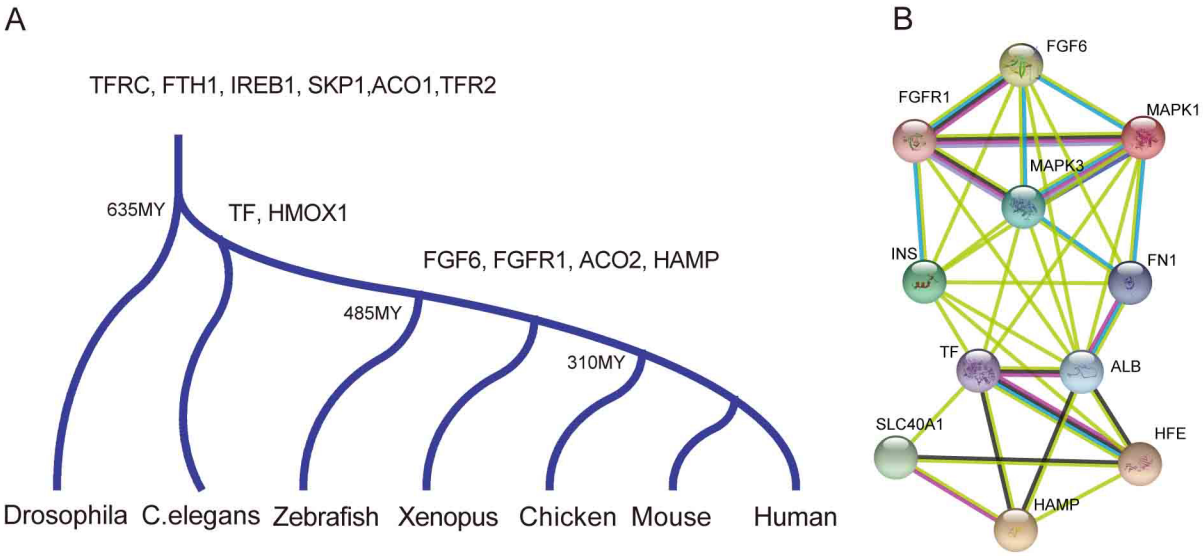

A
TFRC, FTH1, IREB1, SKP1,ACO1,TFR2
635MY
TF, HMOX1
FGF6, FGFR1, ACO2, HAMP
485MY
310MY
Drosophila
C.elegans
Zebrafish
Xenopus
Chicken
Mouse
Human
B
FGF6
FGFR1
MAPK1
MAPK3
INS
FN1
TF
ALB
SLC40A1
HFE
HAMP

Figure S5

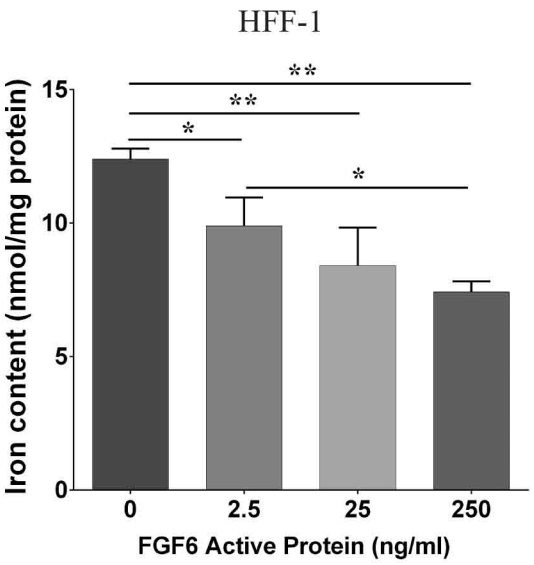

HFF-1
Iron content (nmol/mg protein)
15
10
5
0
**
**
*
*
0
2.5
25
250
FGF6 Active Protein (ng/ml)

Figure S6

FGF6 Mutation Plasmid Structure

M1 Plasmid [GAG -> TAG] **E172X**

ATGGCCCTGGGACAGAAACTGTTCATCACTATGTCCCGGGGAGCA
GGACGTCTGCAGGGCACGCTGTGGGCTCTCGTCTTCCTAGGCATCC
TAGTGGGCATGGTGGTGCCCTCGCCTGCAGGCACCCGTGCCAACA
ACACGCTGCTGGACTCGAGGGGCTGGGGCACCCTGCTGTCCAGGT
CTCGCGCCGGGCTAGCTGGAGAGATTGCCGGGGTGAACTGGGAAA
GTGGCTATTTGGTGGGGATCAAGCGGCAGCGGAGGCTCTACTGCA
ACGTGGGCATCGGCTTTCACCTCCAGGTGCTCCCCGACGGCCGGA
TCAGCGGGACCCACGAGGAGAACCCCTACAGCCTGCTGGAAATTT
CCACTGTGGAGCGAGGCGTGGTGAGTCTCTTTGGAGTGAGAAGTG
CCCTCTTCGTTGCCATGAACAGTAAAGGAAGATTGTACGCAACGCC
CAGCTTCCAAGAAGAATGCAAGTTCAGAGAAACCCTCCTGCCCAA
CAATTACAATGCCTACTAGTCAGACTTGTACCAAGGGACCTACATT
GCCCTGAGCAAATACGGACGGGTAAAGCGGGGCAGCAAGGTGTC
CCCGATCATGACTGTCACTCATTTCCTTCCCAGGATCTAA

Point Mutation Construct Primers

F1 : ATACTCGGATCCGCCACCATG
R1 : CTTGGTACAAGTCTGACTAGTAGGCATTG
F2 : CAATGCCTACTAGTCAGACTTGTACCAAG
R2 : GCTGCAGAATTCTTACGTAATCTGGAA

FGF6 Mutation Plasmid Structure

M2 Plasmid [GAC -> GTC] **D174V**

ATGGCCCTGGGACAGAAACTGTTCATCACTATGTCCCGGGGAGCA
GGACGTCTGCAGGGCACGCTGTGGGCTCTCGTCTTCCTAGGCATCC
TAGTGGGCATGGTGGTGCCCTCGCCTGCAGGCACCCGTGCCAACA
ACACGCTGCTGGACTCGAGGGGCTGGGGCACCCTGCTGTCCAGGT
CTCGCGCCGGGCTAGCTGGAGAGATTGCCGGGGTGAACTGGGAAA
GTGGCTATTTGGTGGGGATCAAGCGGCAGCGGAGGCTCTACTGCA
ACGTGGGCATCGGCTTTCACCTCCAGGTGCTCCCCGACGGCCGGA
TCAGCGGGACCCACGAGGAGAACCCCTACAGCCTGCTGGAAATTT
CCACTGTGGAGCGAGGCGTGGTGAGTCTCTTTGGAGTGAGAAGTG
CCCTCTTCGTTGCCATGAACAGTAAAGGAAGATTGTACGCAACGCC
CAGCTTCCAAGAAGAATGCAAGTTCAGAGAAACCCTCCTGCCCAA
CAATTACAATGCCTACGAGTCAGTCTTGTACCAAGGGACCTACATT
GCCCTGAGCAAATACGGACGGGTAAAGCGGGGCAGCAAGGTGTC
CCCGATCATGACTGTCACTCATTTCCTTCCCAGGATCTAA

Point Mutation Construct Primers

F1 : ATACTCGGATCCGCCACCATG
R1 : GGTACAAGACTGACTCGTAGGCATTG
F2 : CTACGAGTCAGTCTTGTACCAAGGGA
R2 : GCTGCAGAATTCTTACGTAATCTGGAA

FGF6 Mutation Plasmid Structure

M3 Plasmid [CGG ->CAG] **R188Q**

ATGGCCCTGGGACAGAAACTGTTCATCACTATGTCCCGGGGAGCA
GGACGTCTGCAGGGCACGCTGTGGGCTCTCGTCTTCCTAGGCATCC
TAGTGGGCATGGTGGTGCCCTCGCCTGCAGGCACCCGTGCCAACA
ACACGCTGCTGGACTCGAGGGGCTGGGGCACCCTGCTGTCCAGGT
CTCGCGCCGGGCTAGCTGGAGAGATTGCCGGGGTGAACTGGGAAA
GTGGCTATTTGGTGGGGATCAAGCGGCAGCGGAGGCTCTACTGCA
ACGTGGGCATCGGCTTTCACCTCCAGGTGCTCCCCGACGGCCGGA
TCAGCGGGACCCACGAGGAGAACCCCTACAGCCTGCTGGAAATTT
CCACTGTGGAGCGAGGCGTGGTGAGTCTCTTTGGAGTGAGAAGTG
CCCTCTTCGTTGCCATGAACAGTAAAGGAAGATTGTACGCAACGCC
CAGCTTCCAAGAAGAATGCAAGTTCAGAGAAACCCTCCTGCCCAA
CAATTACAATGCCTACGAGTCAGACTTGTACCAAGGGACCTACATT
GCCCTGAGCAAATACGGACAGGTAAAGCGGGGCAGCAAGGTGTC
CCCGATCATGACTGTCACTCATTTCCTTCCCAGGATCTAA

Point Mutation Construct Primers

F1 : ATACTCGGATCCGCCACCATG
R1 : CGCTTTACCTGTCCGTATTTGCTC
F2 : GAGCAAATACGGACAGGTAAAGCG
R2 : GCTGCAGAATTCTTACGTAATCTGGAA

Figure S7

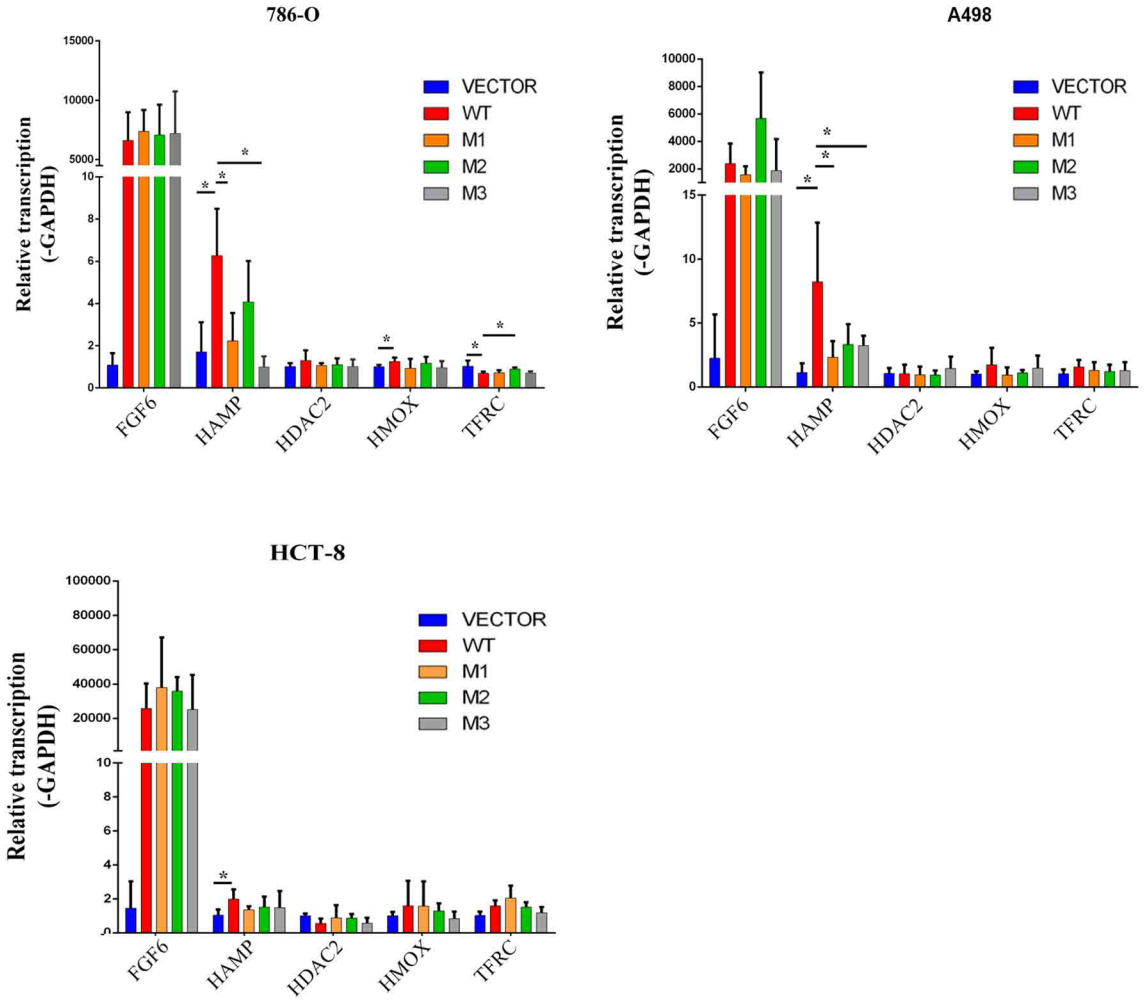

786-O
A498
HCT-8
Relative transcription (-GAPDH)
VECTOR
WT
M1
M2
M3
FGF6
HAMP
HDAC2
HMOX
TFRC

Figure S8

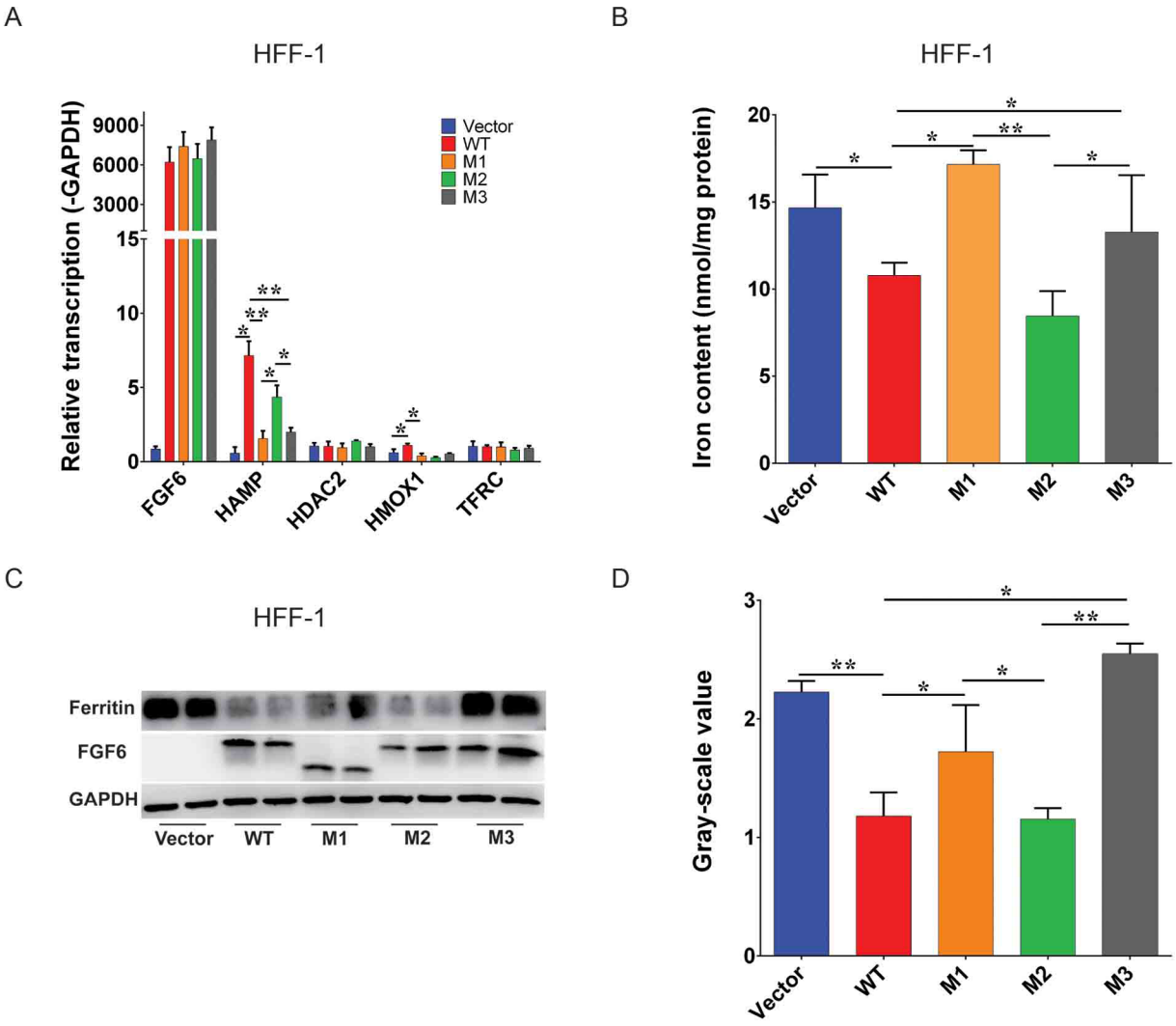

A
HFF-1
Relative transcription (-GAPDH)
9000
6000
3000
15
10
5
0
Vector
WT
M1
M2
M3
FGF6
HAMP
HDAC2
HMOX1
TFRC
B
HFF-1
Iron content (nmol/mg protein)
20
15
10
5
0
Vector
WT
M1
M2
M3
C
HFF-1
Ferritin
FGF6
GAPDH
Vector
WT
M1
M2
M3
D
Gray-scale value
3
2
1
0
Vector
WT
M1
M2
M3

Figure S9

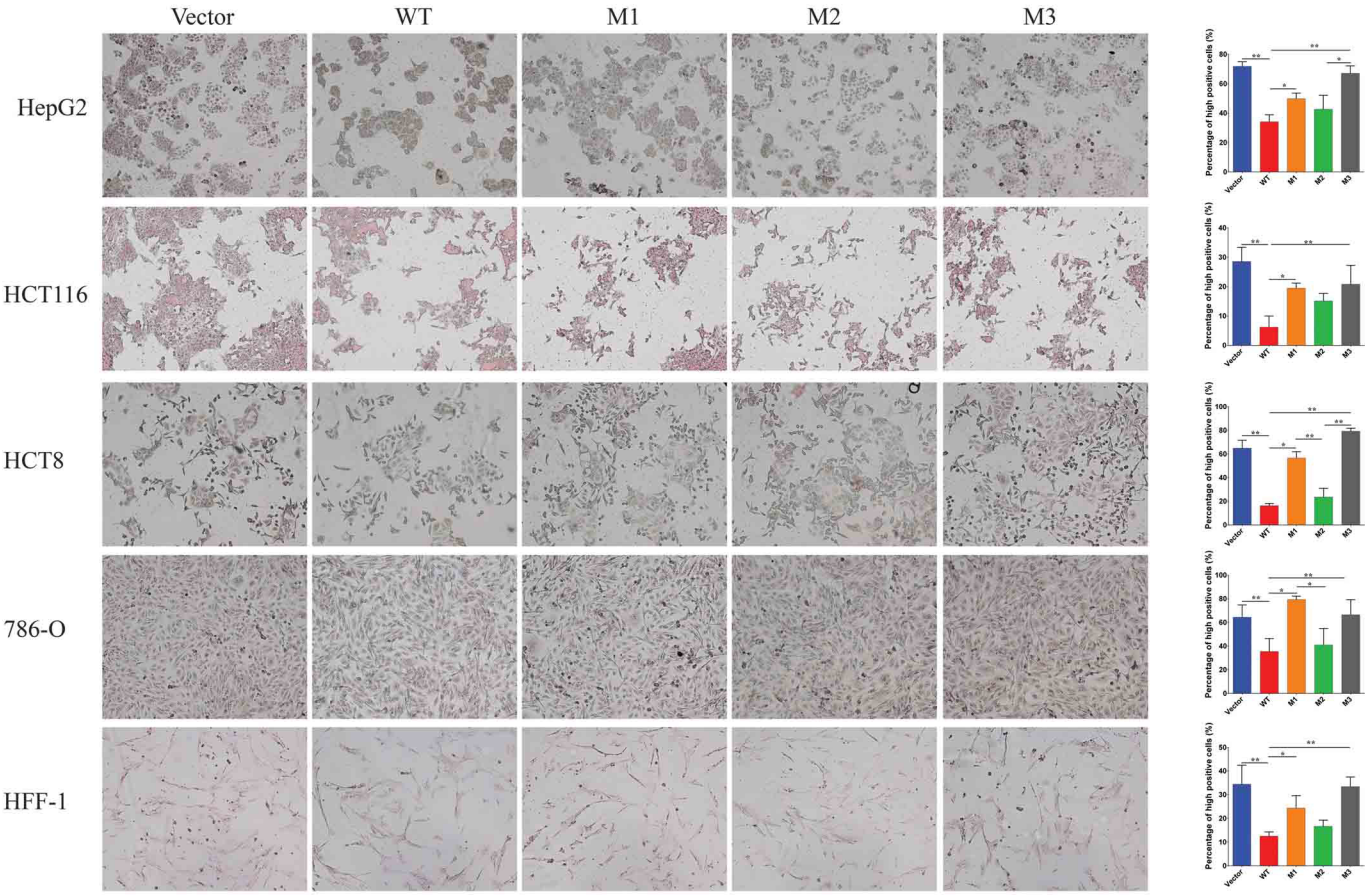

Vector
WT
M1
M2
M3
HepG2
HCT116
HCT8
786-O
HFF-1
Percentage of high positive cells (%)

Figure S10

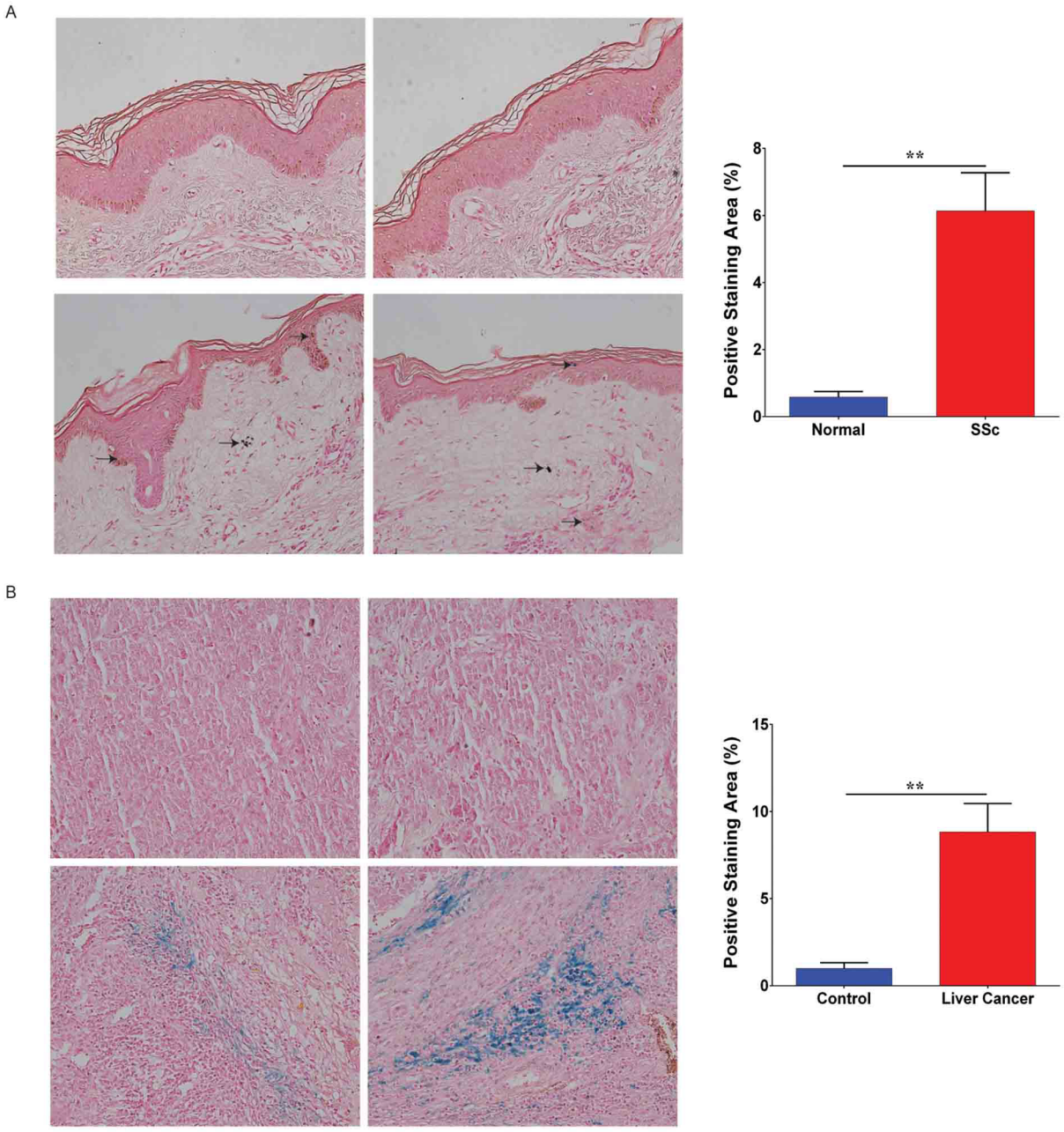

A
Positive Staining Area (%)
8
6
4
2
0
**
Normal
SSc
B
Positive Staining Area (%)
15
10
5
0
**
Control
Liver Cancer

Figure S11

A (normal hepatocytes and metastatic cells) B (non-metastatic liver cancer cells)