Article Title: A Gene-Based Recessive Diplotype Exome Scan Discovers *FGF6*, a Novel Hepcidin-Regulating Iron Metabolism Gene

Short Title: *FGF6* in Iron Metabolism

Shicheng Guo[1#], Shuai Jiang[2#], Narendranath Epperla[3], Yanyun Ma[2], Mehdi Maadooliat[1,4], Zhan Ye[5], Brent Olson[5], Minghua Wang[6], Terrie Kitchner[1], Jeffrey Joyce[1], Robert Strenn[5], Joseph J. Mazza[7], Jennifer K. Meece[8], Wenyu Wu[9], Li Jin[2], Judith A. Smith[10], Jiucun Wang[2,11,*], Steven J. Schrodi[1,12,*]

[1]Center for Human Genetics, Marshfield Clinic Research Institute, Marshfield, WI, USA
[2]State Key Laboratory of Genetic Engineering, Collaborative Innovation Center for Genetics and Development, School of Life Sciences, Fudan University, Shanghai, China
[3]Division of Hematology, The Ohio State University, Columbus, OH, USA
[4]Department of Mathematics, Statistics and Computer Science, Marquette University, Milwaukee, WI, USA
[5]Biomedical Informatics Research Center, Marshfield Clinic Research Institute, Marshfield, WI, USA
[6]Department of Biochemistry and Molecular Biology, Medical College, Soochow University, Suzhou, China
[7]Clinical Research Center, Marshfield Clinic Research Institute, Marshfield, WI, USA
[8]Integrated Research and Development Laboratory, Marshfield Clinic Research Institute, Marshfield, WI, USA
[9]Department of Dermatology, Huashan Hospital, Fudan University, Shanghai, China
[10]Department of Pediatrics, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI, USA
[11]Institute of Rheumatology, Immunology and Allergy, Fudan University, Shanghai, China
[12]Computation and Informatics in Biology and Medicine, University of Wisconsin-Madison, Madison, WI, USA

[#]These authors contributed equally to this work

*Correspondence:
Jiucun Wang, Ph.D.
School of Life Sciences
Fudan University
Shanghai, China
Email: jcwang@fudan.edu.cn

Steven J. Schrodi, Ph.D.
Center for Human Genetics
Marshfield Clinic Research Institute
1000 N Oak Ave--MLR
Marshfield, WI 54449
Tel: 715-221-6443
Email: Schrodi.steven@mcrf.mfldclin.edu
Email: Schrodi@wisc.edu

Abstract

Standard analyses applied to genome-wide association data are well-designed to detect additive effects of moderate strength. However, the power for standard GWAS analyses to identify effects from recessive diplotypes is not typically high. We proposed and conducted a gene-based compound heterozygosity test to reveal additional genes underlying complex diseases. With this approach applied to iron overload, a strong association signal was identified between the fibroblast growth factor-encoding gene, *FGF6,* and hemochromatosis in the central Wisconsin population. Functional validation showed FGF-6 regulates iron homeostasis and induces transcriptional regulation of hepcidin. Moreover, specific identified *FGF6* variants differentially impact iron metabolism. In addition, FGF6 downregulation correlated with iron metabolism dysfunction in systemic sclerosis and cancer cells. Using the recessive diplotype approach revealed a novel susceptibility hemochromatosis gene and has extended our understanding of the mechanisms involved in iron metabolism.

Key Points

- An exome scan for recessive effects reveals *FGF6* as a hemochromatosis susceptibility gene.
- FGF-6 decreases ferrous iron uptake in liver cells and induces increased hepcidin expression.

Introduction

Genome-wide Association Studies (GWAS) are well-designed to detect additive effects of modest effect sizes. We hypothesized that gene-based tests sensitive to recessive diplotypes—including recessive single site effects and compound heterozygosity—may reveal additional genes underlying complex diseases. Carrying variants conferring a compromised function on both homologous chromosomes is likely to impact molecular physiological states. Deep sequencing studies have conclusively shown a vast reservoir of rare variants segregating in human populations.[1] Rare variants in functional categories (e.g., missense, regulatory motifs) may generate pathogenic effects through recessively-acting diplotypes, and such effects are apt to remain concealed from standard GWAS analyses. Simple power calculations show that recessive diplotype inheritance produces signals that are difficult for standard GWAS methods to discover (**Fig. S1**). Compound heterozygosity disease models also enjoy a high degree of biological plausibility, particularly if the alleles confer compromised protein function.[2-5] Recessive diplotype modes of inheritance are well-established in Mendelian diseases, such as cystic fibrosis[6], mevalonic aciduria[7], beta-thalassemia[8] and Niemann-Pick disease.[9] Although not systematically examined in population-based studies, there is a sizable repository of genes underlying complex diseases with recessive, loss-of-function effects.[10-14] Hence, we posited that an exome-wide, gene-based screen of recessive diplotypes using putative functional variants in both oligogenic and complex diseases may expand our knowledge of disease genes.

Iron metabolism disorders, including adult hereditary hemochromatosis, collectively are common conditions with considerable public health implications.[15,16] Importantly, the hepatic hormone hepcidin is a key regulator of iron homeostasis by controlling iron flux from enterocytes and macrophages to plasma through degradation of the cellular iron exporter ferroportin (*SLC40A1*). Within cells, ferritin is the iron-storage protein which can be used for indirect iron quantification. To investigate the inheritance of hemochromatosis, several segregation analyses were initially conducted, concluding that a recessive mode of inheritance is highly plausible.[17,18] Several human studies have investigated the genetics of iron overload, revealing several critically important genes. Notably, *HFE*, encoding the membrane-bound hereditary hemochromatosis protein, was mapped two decades ago through family-based linkage[19-22] and association approaches.[23,24] Additional studies have definitively placed the missense polymorphism C282Y (rs1800562) in *HFE* as the major susceptibility factor in adult-onset, type 1 hereditary hemochromatosis.[25,26] Additional genes have been identified through pathway-based genetic association studies and GWAS, including *BMP2, BMP4, HJV, TF, TMPRSS6, NAT2, FADS2*, and *TFR2*.[27-29]

Methods

Central Wisconsin Hemochromatosis Sample Set

The homogenous population in rural Central Wisconsin is the source population for the Personalized Medicine Research Project (PMRP), a biobank linked to electronic health records (EHRs) housed by the Marshfield Clinic Research Institute.[30] Samples from over 20,000 individuals comprise the PMRP. The study was conducted in accordance with the Declaration of Helsinki. All samples were collected following written informed consent. All investigators using the PMRP samples had obtained Research Ethics and Compliance Training certification through the CITI program. The study protocol was reviewed and approved by the Marshfield Clinic Institutional Review Board (details in **Acknowledgements**). The Central Wisconsin population is largely stationary and primarily derived from Bavarian migrants in the late 1800s. The population carries high utility for disease gene mapping through reduction in confounding by population stratification and lower expected levels of allelic and locus heterogeneity. Additionally, environmental exposures are thought to be relatively uniform across this population. For these reasons, the PMRP has been effectively used in numerous human genetics studies.[31-34] PMRP DNA samples were collected and stored approximately 14 years ago and all individuals have longitudinal EHR information housed at the Marshfield Clinic, averaging >30 years. The EHR is composed of ICD-9 diagnostic codes, laboratory test results, clinical procedure data, prescription information and physician notes. Hemochromatosis cases and controls were selected from the PMRP population. Hemochromatosis cases were selected on the basis of percent transferrin saturation laboratory values (the ratio of serum iron to transferrin iron-binding capacity) exceeding 48% and having two or more instances of ICD-9 codes indicating the diagnosis of hemochromatosis: 275.0 (iron metabolism disorder, excluding anemia), 275.01 (hereditary hemochromatosis), 275.03 (unspecified hemochromatosis), and/or 275.09 (other iron metabolism disorders). To reduce confounding by population stratification, a Principal Components Analysis (PCA) on the exome genotyping data was implemented using all samples, blinded to disease status. Individuals considered genetic background outliers (more than three standard deviations from the centroid of the first two principal components) were excluded from the study. Following the removal of outliers, the resulting set of individuals was highly homogeneous based on the first three principal components. Exhaustive pairwise kinship coefficients were calculated and one individual from pairs of individuals exhibiting third-degree or closer relatedness were removed. Of the approximately 10,000 individuals previously subjected to the exome genotyping array and quality control procedures, the phenotype algorithm identified 18 individuals that were selected as hemochromatosis cases. Controls (n=6,896) were individuals without abnormal saturation values and without any instances of hemochromatosis ICD-9 codes.

Genotyping

Of the full PMRP cohort, approximately 10,000 DNAs were interrogated by high density genotyping on the Illumina HumanCoreExome beadchip. This exome array of >500,000 markers has ancestry informative markers, a panel of identity-by-descent SNPs, coverage of markers found to be genome-wide significant in

GWAS studies, and excellent coverage of exonic variants. The version of the beadchip was designed and used in the AMD consortium.[34] Rare variants (<1% frequency) represented 47.8% of the markers, moderately common variants (1-10% frequency) were 8.1% of the variants, and 44.1% of the variants interrogated were common alleles (>10% frequency). The genotyping quality control measures were previously described (call rates for each variant or individual >0.985).[34] Variants exhibiting departure from Hardy-Weinberg equilibrium ($p<1x10^{-6}$) were excluded from subsequent analyses. Additional recent studies have used data generated from this genotyping to discover susceptibility genes for common diseases.[35] Following QC procedures, 413,701 variants remained for analysis. The site frequency spectrum of the resulting variants is displayed in **Fig. S2** and **Supplementary Table 1.**

Haplotype Phasing

In general, gametic phasing is necessary to directly determine compound heterozygous individuals at a particular gene. Using all 10,000 exome-genotyped samples from the PMRP, the software package Beagle was applied to infer phased haplotypes from unphased genotype data using a localized haplotype-cluster model algorithm.[36] The calculations were performed on a high performance computing cluster housed at the Marshfield Clinic. As the subsequent analyses were gene-based and the genotyping data was concentrated on exonic variants, each gene in the exome was phased separately using this approach. Although rare variants can present difficulties in phasing, the use of a large sample size from a highly homogeneous population aids in mitigating the error rate. Notably, Beagle has been shown to have error rates in phasing between 0.77-0.94% for medium (n=1,000) to large (n=5,000) sample sizes using a 500K GWAS array.[36] Recently, the switch error rate was calculated for the Beagle applied to two sequencing datasets. Beagle attained a switch error rate of 1.525% and 0.488% for the 1000 Genomes Project and Haplotype Reference Consortium, respectively.[37]

Determination of Putative Functional Variants

Following the phasing of the genotype data, putative functional variants were identified. The putative functional variants included in the analyses satisfied the quality control criteria as described previously.[34] Markers used in the analyses were either GWAS-significant as of June 2015 and/or annotated as missense, nonsense, 3'UTR, 5'UTR or occurring within a splice site region. Additionally, on the resulting set of variants, annotation was performed using the ANNOVAR software.[38] Only variants that were also annotated as pathogenic by at least two ANNOVAR predictions were included in the scan. Following filtering for putative functional variants, 129,556 SNPs remained for use in our gene-based recessive diplotypes scan. The putative functional variants of *FGF6* are shown in **Supplemental Table 2.**

Statistical Tests of Recessive Diplotypes

At each gene, individuals were classified as having a recessive diplotype configuration if they carried at least one putative functional allele on each homolog (PF). Individuals carrying at least one homolog free from

putative functional alleles were deemed as having a wildtype diplotype (W). The total number of case/control individuals carrying a recessive diplotype was denoted by PF_{cs} and PF_{ct}, respectively. Similarly, the total number of case/control individuals carrying a wildtype diplotype was denoted by W_{cs} and W_{ct}, respectively. Following the determination of these counts, a Fisher's exact test was applied. As the hypergeometric null density holds for all sample sizes, the Fisher's exact test is robust to imbalance between case and control sample sizes. Simulations have recapitulated this finding showing that the Fisher's exact test does not inflate type I error rates under unbalanced designs.[39] Individuals carrying one or more homozygous genotype(s) at a single site for a putative functional allele were included in the PF category. Genes without any high quality, putative functional alleles across all samples were removed from the analyses. Across all genes with analyzable data, a conservative experiment-wise multiplicity correction was calculated using 15,900 gene-based tests. To compare the recessive diplotype analysis procedure to a standard rare variant gene-based test, the RVTESTS software[40] which implements the sequence kernel association test was also applied to the genotype data.[41] Additionally, to investigate the sex-specific effects, the Haldane OR was calculated separately for female and male strata. Lastly, the Mantel-Haenszel test of homogeneity was calculated to determine the level of statistical evidence for sex-specific differences in effects.[42]

Power Calculations

To explore the efficacy of the proposed approach, we performed analytic power calculations under the alternative model of compound heterozygosity/recessive inheritance of disease at two sites, each segregating two alleles. By doing so, we sought to compare the power of a standard GWAS analysis (Armitage trend test) to a log-likelihood ratio G test.[42] **Fig. S1** shows the power of each of these tests across different sets of penetrances and haplotype frequencies. To consolidate the different sets of haplotype frequencies, the results are plotted as a function of linkage disequilibrium between the two sites. The power of the G test for recessive diplotypes exceeded the power for the Armitage trend test across virtually all of the parameter space. Additional work in this area was recently performed by Sanjak and colleagues showing similar results.[43]

Comparative genomic analysis and Protein-protein interaction inference

Amino acid sequencing of the core iron metabolism genes were collected including *TFRC, FTH1, IREB1, SKP1, SKP1, ACO1, TFR2, TF, HMOX1, ACO2, HAMP, FGF6, and FGFR1*. The alignments were derived from NCBI BLASTn database. Phylogeny for different genes were compared to show the earliest evolutionary time point and then occurrence for each gene were mapped to the phylogenetic tree.[44] Protein-protein interaction network inference was conducted to FGF-6 and main iron metabolism proteins. The finial network was tuned after removing non-necessary nodes between FGF-6 and key iron molecules including HFE and SLC40A1.

Cell culture, Reagents, and Protein treatment

Colon cancer cell lines (HCT8 and HCT116), a kidney cancer cell line (786-O), a liver cancer cell line (HepG2) and a fibroblast cell line (HFF-1) were cultured in DMEM medium supplemented with 10% FBS at 37°C in a 5% CO2 humidified incubator. To investigate the change in iron uptake under different protein treatments or plasmid transfection, the Ferric Ammonium Citrate (FAC) cell culture method was employed. Cells were cultured in normal DMEM and FBS medium with the presence of 10uM FAC and 500uM ascorbate for 48h during detection of cellular iron concentration. Cells cultured in normal medium exhibited very low iron concentrations. Total intracellular iron concentration in cells were considerably higher than cells Recombinant human FGF-6 protein (Active) (ab219122), anti-Ferritin (ab75973) were purchased from Abcam, Flag tag antibody (20543-1-AP) was purchased from Proteintech Group, anti-GAPDH antibody was purchased from Shanghai Yeasen Biotechnology. 1mM Ferric ammonium citrate and 50mM ascorbate were dissolved in distilled water. NaOH, HCl, KMnO4, ferrozine, neocuproine, ammonium acetate, ascorbic acid and FeCl3 were purchased from Beijing Oka Biological Technology. Plasmid with raw *FGF6* sequence was purchased from PPL-Shanghai Co., Ltd (Shanghai, China) which was constructed in an N-Terminal p3XFLAG-CMV vector, whereas three different *FGF6* mutations (E127X, D174V, R188Q) were synthesized with overlapping-PCR.

Quantification of Iron content by ferrozine assay

Total intracellular iron content was measured by the ferrozine assay.[45] Cells were cultured in 12 or 24-well plate for 48h and washed three times with cold PBS. After lysed 2h with 50mM NaOH, 100μL of cell lysates were mixed with 10mM HCl, and 100μL of the iron-releasing reagent (a freshly mixed solution of equal volumes of 1.4M HCl and 4.5% (w/v) KMnO4 in H2O). The mixtures were incubated for 2h and 30μL iron detection reagent (6.5 mM ferrozine, 6.5 mM neocuproine, 2.5M ammonium acetate, and 1M ascorbic acid) was added, after 30min incubation, 280μL of solution was added to a 96-well plate and read 550nm on a microplate reader. In addition, FeCl3 (0-100μM) was used as iron standards and protein quantification was determined by a Lowry protein assay.

Western blot

Cell lysates were harvested when incubated iron for 48h, then equal amounts of protein from every sample were subjected to 12% SDS-PAGE gels electrophoresis and then transferred to PVDF membranes. After blocked with 5% BSA, the membranes were incubated with GAPDH (1:10000), Ferritin (1:1000) and Flag (1:2000) at 4℃ overnight. Then, membranes were washed three times with TBST, incubated with anti-rabbit or anti-mouse secondary antibody. The bands were visualized using Image QuantTL software.

Perls' staining

Cells were washed with phosphate-buffered saline (PBS) three times, fixed with 4% glutaraldehyde for 10 min, and incubated at 37°C for 60 min with 2 ml Prussian blue solution comprising equal volumes of 2% hydrochloric acid aqueous solution and 2% potassium ferrocyanide (II) trihydrate. After the cells were stained

with 0.5% neutral red for 3 min, iron staining was visualized by Nikon microscope. Iron positive high positive staining cells divided by total cell number used to evaluate the iron deposition levels.

RT-PCR and quantitative RT-PCR analysis

Total RNA was extracted from the cells using Trizol (Invitrogen, Carlsbad, CA, USA). One microgram of total RNA was subjected to cDNA synthesis using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) according to the manufacturer's instructions. The specific primers for each gene were designed using Primer 5 and synthesized by Generay Biotech Co., Ltd. (Shanghai, China). The RT-PCR amplification was conducted using a SYBR Green I PCR Kit (TaKaRa, Shanghai, Japan) according to manufacturer's instructions. The reaction was carried out on a ABI Prism 7900 Detector System (Applied Biosystems). RT-PCR conditions were 95°C for 3 min, followed by 40 cycles of 95°C for 15 s, 60°C for 40 s, and the conditions for obtaining the dissociation curve were 95°C for 15 s, 60°C for 15 s, 95°C for 15 s. The data obtained from the assays were analyzed with SDS 2.3 software (Applied Biosystems). For each sample, the relative gene expression was calculated using a relative ratio to GAPDH. RT-PCR primers can be found in the **Supplementary Table 3**.

Immunohistochemical staining of FGF-6

The primary antibody used was anti-FGF-6 (1:200, D162668 BBI, Shanghai) and anti-Ferritin (1:100, ab75973, abcam, China). Liver and skin tissues from four liver cancer patients and six SSc patients, respectively, and normal controls were formalin-fixed and paraffin-embedded. Sections were deparaffinized and incubated with 5% bovine serum albumin for 60 min. Cells positive for FGF-6 were detected by incubation with the primary antibody for 2 h at room temperature followed by incubation with 3% hydrogen peroxide for 10 min. Rabbit anti-rabbit IgG labeled with horseradish peroxidase were used as secondary antibodies. The expression of FGF-6 or ferritin was visualized with 3,3-diaminobenzidinetetrahydrochloride (DAB-4HCl). The expression of FGF-6 in SSc and tumor tissues was quantitated by the average optical density (AOD) of positive signal in each sample using the software imageJ (Windows and Java-1.8.0, NIH).

Results

Gene-based compound heterozygosity scan identifies a novel hemochromatosis susceptibility gene

To discover novel iron overload-predisposing genes, we conducted a gene-based scan for recessive diplotypes composed of putative functional alleles across the exome using biobanked samples linked to electronic medical records obtained from a rural, genetically-homogeneous population in central Wisconsin. Of the 10,000 samples evaluated, our transferrin saturation and diagnostic code-based phenotype algorithm identified 18 case individuals and 6896 controls. We estimated gametic phase on all individuals and restricted our analyses of diplotypes to putative functional variants. Our recessive diplotype scan identified two exome-

wide significant genes (**Figure 1, Table 1 and Fig. S3**), *HFE* (P=1.29×10^{-8}; OR=28.7) and *FGF6 (P=1.99*×10^{-6}; OR=22.8*)*. For comparison, the SKAT/rvtest procedure on the *FGF6* genotype data yielded an asymptotic *P=3.86*×10^{-5} and permuted *P=1.0*×10^{-4}. Notably, the recessive diplotype scan result exceeded exome-wide significance, whereas the rare variant test did not. Additionally, no statistical evidence of effect differences between females and males for the *FGF6* data (Mantel-Haenszel test of homogeneity P=0.728). These results motivated our investigation of FGF-6 function and the impact of specific *FGF6* variants on iron metabolism.

Protein-protein interaction indicates FGF-6 is involved in iron metabolism network

To explore the involvement of *FGF6* in iron metabolism, we found evidence for FGF-6 interactions with FGFR1, MAPK1/3, INS, FN1 and involvement in the iron metabolism subnetwork involving TF, HFE, HAMP and SLC40A1 (**Fig. S4**) by investigating FGF-6 protein-protein interactions. FGF-6, also known as Heparin Secretory-Transforming Protein 2 or Heparin-Binding Growth Factor 6 has multiple heparin binding sites (HBS). Three known nonsynomymous variants located in the heparin binding sites (R188Q) or flanking sites (D174V and E172X) were speculated to be important for the FGF-6 function. Further, D174V and E172X are located in the regions between FGFR-binding region (FGFR-BR-3) and HBS-1 (**Figure 2**). Hence, we studied these three variants in functional studies to further investigate the involvement of FGF-6 in iron metabolism.

FGF-6 modulation of hepcidin expression and iron uptake

To investigate the potential mechanism linking FGF-6 to iron metabolism, the effects of FGF6 on iron uptake and the expression of iron-metabolism genes in HepG2, HCT8, HCT116, 786-O and HFF1 cells were evaluated. Using cells cultured cells and a ferrozine assay to detect iron, total intracellular iron concentration was significantly decreased in HepG2, 786-O, HCT8, HCT116 and HFF-1 cells when treating with active FGF-6 protein in a dose-dependent manner (**Figure 3A-D and Fig. S5**). Testing the effect of FGF6 on the expression of a set of genes involved in iron metabolism (*HAMP*, *HDAC2*, *HMOX1*, *TFRC*, and *HEPH*), HepG2 cells were subjected to treatment from control or FGF-6 protein and *FGF6* mRNA or control and mRNA expression relative to *GAPDH* was measured in the five iron metabolism genes. RT-PCR analysis revealed that *HAMP* and *HDAC2* mRNA levels were significantly increased after the FGF-6 active protein introduction in HepG2 cells compared to treatment with PBS as control (**Figure 4A**). *FGF6* plasmid transfection significantly increased *HAMP*, *HDAC2*, and *HMOX1* levels, whereas *TFRC* levels significantly decreased in HepG2 compared to a vector without *FGF6* (**Figure 4B**). *HEPH* expression did not change with either *FGF6* plasmid or FGF-6 protein, suggesting the effect of FGF-6 may be independent of *HEPH* (**Figures 4A, 4B**).

Evaluation of *FGF6* variants on HAMP expression and iron concentration compared to wildtype *FGF6*

In order to investigate the effects of the *FGF6* alleles on FGF-6 function, we transfected plasmids carrying either the wildtype *FGF6* or variant *FGF6* with each of the three point mutations described above (**Figure 2 and Fig. S6**). The M1 (E172X) and M3 (R188Q) variants exhibited a significant downregulation of *HAMP* mRNA compared to WT in HepG2 cells (**Figure 4C**), HCT-116 cells (**Figure 4D**), 786-O (**Fig. S7**), A498 (**Fig.**

S7), but not HCT-8 cells (Fig. S7). Evaluating the effect of M2 (D174V) on *HAMP* expression compared to WT only yielded a significant reduction in HepG2 (**Figure 4C**), but not in any of the other cell lines (**Figure 4D**, Fig. S7). Further, we noted *HAMP* mRNA levels in M1 and M3 transfections were comparable to control levels, which illustrated a strong attenuation of *FGF-6* function for M1 and M3 variants (**Figure 4C-D,** Fig. S7). Examining the impact of specific variants on intracellular iron concentration in HepG2 and HCT-116 cells, M1 and M3 produced significantly elevated iron deposition **(Figure 4E, 4F,** Fig. S8**)** and ferritin expression (**Figure 4G-H,** Fig. S8) indicating a deficiency in M1/M3 *FGF6*-mediated iron homeostasis. In addition, the intracellular iron accumulation pattern was confirmed by IHC using Perls' stain **(**Fig. S9**)**. In contrast, M2 did not produce a significant departure from WT in iron concentration and ferritin expression in HepG2 and HCT-116 (**Figure 4**). Further, TFRC expression was significantly upregulated in the presence of M3 compared to WT (**Figure 4C**). The functions mentioned above were also validated in HFF-1 (Fig. S8).

Altered *FGF6* gene expression in systemic sclerosis and cancer

We hypothesized that FGF-6 might be involved in human autoimmune diseases and cancers since abnormal iron metabolism has been reported in numerous studies.[47-50] More specifically, decreased hepcidin has been implicated in the anemia of chronic disease which frequently accompanies these systemic inflammatory states. To explore the relationship between *FGF6* expression and iron deposition in autoimmune tissues, *FGF6* expression and iron deposition in the skin lesions from systemic sclerosis patients (SSc) and healthy controls were examined. We found significantly decreased FGF-6 protein by immunohistochemistry assay (**Figure 5A**) and elevated iron deposition in SSc skin tissue by Ferrozine assay (**Figure 5B**), especially in the epidermis. Increased iron deposition was confirmed by Perl's stain in SSc skin tissues (Fig. S10A). We also investigated the relationship between FGF6 protein expression with iron deposition in liver cancer tissues. We found FGF-6 was significantly decreased in non-metastatic cancer lesion tissues **(Figure 5C)** and the increased iron deposition (**Figure 5D,** Fig. S10B). However, increased FGF-6 expression was observed in metastatic liver carcinoma tissue (Fig. S11), suggesting that FGF-6 plays different roles in oncogenesis and metastasis, analogous to TGF-β.[51,52]

Discussion

Iron homeostasis results from a combination of pathways and four main cell types: enterocyte, hepatocyte, macrophage, and erythroblast. The EGF/EGFR signaling pathway, heme production, STAT signaling, cAMP signaling, ferritin storage, and BMP-SMAD signaling are all involved in iron regulation. We conducted an exome-wide, gene-based recessive diplotype scan using putative functional variants to reveal additional genes underlying hemochromatosis susceptibility—an approach that can be widely applied to investigate complex disease susceptibility generated by compound heterozygosity and recessive single site effects using existing exome-wide association genotype and sequencing data. Although the case sample size was very small, this novel scan identified *FGF6* as being significantly associated with hemochromatosis following correction for

multiple testing. *FGF6* belongs to the paracrine FGF-gene family and is largely expressed in skeletal muscle, which plays an important role in iron metabolism as it contains 10%-15% of iron stores. We conducted the evolutionary analysis of *FGF6* and known iron metabolism genes including *FGFR1, TFRC, FTH1, IREB1, TF, HMOX1, ACO2* and *HAMP* (encoding hepcidin). The appearance of iron metabolism genes can be separated into two stages. *TF* and *HMOX1*, which are found in animals from *C. elegans* to *H. sapiens*, indicate an origin in early Bilateria evolution (~635 Mya). *FGF6*, *FGFR1*, *ACO2* and *HAMP* can be found from *D. rerio* to *H. sapiens*, but are not present in *C. elegans* and *Drosophila*, indicating emergence in early Vertebrata (~485 Mya). The co-appearance of these genes suggests possible co-regulatory functions (**Fig. S4A**). Functional experiments demonstrated that FGF-6 strongly impacted hepcidin expression, thereby playing a role in regulation of iron homeostasis. These results suggest FGF-6 mediates its effect on iron metabolism via hepcidin. The induction of hepcidin expression by FGF-6 leads to degradation of ferroportin through binding and internalization of ferroportin by hepcidin as shown in **Figure 6**. We additionally found that three *FGF6* nonsynonymous variants increased intracellular iron concentrations and reduced hepcidin levels compared to wildtype *FGF6*, indicating loss-of-function. Interestingly, a genome-wide RNAi profiling study reported that knockdown of *FGF6* increased transferrin-mediated endocytosis.[53] Rs12368351, approximately 8kb downstream of *FGF6* has been associated with phosphorus levels;[54] and two SNPs, rs140668749 and rs10849061, within 20kb downstream of *FGF6*, are associated with migraine.[55,56] Previous studies have indicated that iron plays a role in autoimmunity and a study examining pulmonary arterial hypertension in SSc noted iron deposition in lung elastin fibers and giant cells[57], however, epidermal iron deposition in SSc has not been previously investigated. We observed that FGF-6 is involved with iron deposition in SSc and liver cancer. Together, these results demonstrate that fibroblast growth factor receptor (FGFRs) signaling through FGF-6 is a critically important mechanism in iron metabolism.

Acknowledgements

On April 15, 2014, the study was reviewed and approved by the Marshfield Clinic Research Institute Institutional Review Board. FWA00000873, IRB00000673, Title: Two Allele Loss of Function Genotype Array Study. Marshfield Clinic received a Certificate of Confidentiality from the National Institutes of Health. The project described was supported by the Clinical and Translational Science Award (CTSA) program, previously through the National Center for Research Resources (NCRR) grant 1UL1RR025011 and the National Center for Advancing Translational Sciences (NCATS) grant 9U54TR000021, and now by the NCATS grant UL1TR000427. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. Additional funding was provided by the Marshfield Clinic Research Institute grant SCH10218 and generous donors to the Marshfield Clinic.

Authorship Contributions

SG performed analyses, interpreted results, designed the functional experiments, and aided in drafting the manuscript. SJ conducted molecular and cell biology experiments. NE interpreted results, provided hematology and pathway expertise, reviewed and edited the manuscript. MM aided in the analyses and reviewed the manuscript. MW, YM and WW provided clinical and biochemistry advice and aided in drafting the manuscript. MM aided with the experimental design and analyses. ZY performed initial genetic and statistical analyses, performed data management and reviewed the manuscript. BO implemented and refined the phenotyping algorithms. TK and JJ aided in the regulatory paperwork and reviewed the manuscript. RS performed data management tasks. JJM provided clinical advice and reviewed the manuscript. JKM supervised the management of biological samples for genotyping and reviewed the manuscript. LJ reviewed the manuscript and provided general scientific advice. JAS provided molecular and cellular biology advice, clinical advice, reviewed and edited the manuscript. JW supervised the functional experiments, reviewed the manuscript and provided biological advice. SJS designed the study, supervised the genetic analyses, developed phenotyping algorithms, developed analysis methods and power calculations, interpreted results and aided in drafting and editing the manuscript.

Disclosure of Conflicts of Interest

The authors declare no conflict of interest.

Abbreviations

GWAS: Genome-wide Association Studies
EHR: Electronic health record
PMRP: Personalized Medicine Research Project
PCA: Principal Components Analysis
AOD: Average optical density
FAC: Ferric ammonium citrate
HCT8: Human ileocecal colorectal adenocarcinoma cell line
HCT116: Human colon carcinoma cell line
786-O: Human kidney adenocarcinoma cell line
HepG2: Human liver hepatocellular carcinoma cell line
HFF-1: Human skin fibroblast cell line
HBS: Heparin binding sites
HAMP: Hepcidin antimicrobial peptide
HEPH: Hephaestin
FGF6: Fibroblast Growth Factor 6 gene
FGF-6: Fibroblast Growth Factor 6 protein

M1: E172X *FGF6* variant (NM_020996.2:c.514G>T)

M2: D174V *FGF6* variant (NM_020996.2:c.521A>T)

M3: R188Q *FGF6* variant (NM_020996.2:c.563G>A)

SSc: Systemic sclerosis

References

1. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285-91 (2016).
2. MacArthur, D.G. *et al.* A systematic survey of loss-of-function variants in human protein-coding genes. *Science* **335**, 823-8 (2012).
3. Zou, J. *et al.* Quantifying unobserved protein-coding variants in human populations provides a roadmap for large-scale sequencing projects. *Nat Commun* **7**, 13293 (2016).
4. Cohen, J.C. *et al.* Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science* **305**, 869-72 (2004).
5. Andreoletti, G. *et al.* Exome Analysis of Rare and Common Variants within the NOD Signaling Pathway. *Sci Rep* **7**, 46454 (2017).
6. De Braekeleer, M., Allard, C., Leblanc, J.P., Simard, F. & Aubin, G. Genotype-phenotype correlation in cystic fibrosis patients compound heterozygous for the A455E mutation. *Hum Genet* **101**, 208-11 (1997).
7. Prietsch, V. *et al.* Mevalonate kinase deficiency: enlarging the clinical and biochemical spectrum. *Pediatrics* **111**, 258-61 (2003).
8. Thein, S.L. Genetic modifiers of beta-thalassemia. *Haematologica* **90**, 649-60 (2005).
9. Bauer, P. *et al.* NPC1: Complete genomic sequence, mutation analysis, and characterization of haplotypes. *Hum Mutat* **19**, 30-8 (2002).
10. Singh, T. *et al.* Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders. *Nat Neurosci* **19**, 571-7 (2016).
11. Adam, R. *et al.* Exome Sequencing Identifies Biallelic MSH3 Germline Mutations as a Recessive Subtype of Colorectal Adenomatous Polyposis. *Am J Hum Genet* **99**, 337-51 (2016).
12. Hague, S. *et al.* Early-onset Parkinson's disease caused by a compound heterozygous DJ-1 mutation. *Ann Neurol* **54**, 271-4 (2003).
13. Onoufriadis, A. *et al.* Mutations in IL36RN/IL1F5 are associated with the severe episodic inflammatory skin disease known as generalized pustular psoriasis. *Am J Hum Genet* **89**, 432-7 (2011).
14. Dewey, F.E. *et al.* Distribution and clinical impact of functional variants in 50,726 whole-exome sequences from the DiscovEHR study. *Science* **354**(2016).
15. Adams, P.C. & Barton, J.C. Haemochromatosis. *Lancet* **370**, 1855-60 (2007).
16. Andrews, N.C. & Schmidt, P.J. Iron homeostasis. *Annu Rev Physiol* **69**, 69-85 (2007).
17. Saddi, R. & Feingold, J. Idiopathic haemochromatosis: an autosomal recessive disease. *Clin Genet* **5**, 234-41 (1974).
18. Borecki, I.B., Rao, D.C., Yaouanq, J. & Lalouel, J.M. Segregation of genetic hemochromatosis indexed by latent capacity of transferrin. *Am J Hum Genet* **45**, 465-70 (1989).
19. Simon, M., Alexandre, J.L., Bourel, M., Le Marec, B. & Scordia, C. Heredity of idiopathic haemochromatosis: a study of 106 families. *Clin Genet* **11**, 327-41 (1977).
20. Cartwright, G.E. *et al.* Inheritance of hemochromatosis: linkage to HLA. *Trans Assoc Am Physicians* **91**, 273-81 (1978).
21. Edwards, C.Q., Griffen, L.M., Dadone, M.M., Skolnick, M.H. & Kushner, J.P. Mapping the locus for hereditary hemochromatosis: localization between HLA-B and HLA-A. *Am J Hum Genet* **38**, 805-11 (1986).
22. Jazwinska, E.C., Lee, S.C., Webb, S.I., Halliday, J.W. & Powell, L.W. Localization of the hemochromatosis gene close to D6S105. *Am J Hum Genet* **53**, 347-52 (1993).
23. Feder, J.N. *et al.* A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nat Genet* **13**, 399-408 (1996).
24. Jazwinska, E.C. *et al.* Haemochromatosis and HLA-H. *Nat Genet* **14**, 249-51 (1996).

25. Griffiths, W. & Cox, T. Haemochromatosis: novel gene discovery and the molecular pathophysiology of iron metabolism. *Hum Mol Genet* **9**, 2377-82 (2000).
26. Allen, K.J. *et al.* Iron-overload-related disease in HFE hereditary hemochromatosis. *N Engl J Med* **358**, 221-30 (2008).
27. Milet, J. *et al.* Common variants in the BMP2, BMP4, and HJV genes of the hepcidin regulation pathway modulate HFE hemochromatosis penetrance. *Am J Hum Genet* **81**, 799-807 (2007).
28. Benyamin, B. *et al.* Novel loci affecting iron homeostasis and their effects in individuals at risk for hemochromatosis. *Nat Commun* **5**, 4926 (2014).
29. de Tayrac, M. *et al.* Genome-wide association study identifies TF as a significant modifier gene of iron metabolism in HFE hemochromatosis. *J Hepatol* **62**, 664-72 (2015).
30. McCarty, C.A. *et al.* Marshfield Clinic Personalized Medicine Research Project (PMRP): design, methods and recruitment for a large population-based biobank. *Per Med* **2**, 49-79 (2005).
31. Carter, T.C. *et al.* Validation of a metabolite panel for early diagnosis of type 2 diabetes. *Metabolism* **65**, 1399-408 (2016).
32. Ye, Z. *et al.* Genome wide association study of SNP-, gene-, and pathway-based approaches to identify genes influencing susceptibility to Staphylococcus aureus infections. *Front Genet* **5**, 125 (2014).
33. Ritchie, M.D. *et al.* Genome- and phenome-wide analyses of cardiac conduction identifies markers of arrhythmia risk. *Circulation* **127**, 1377-85 (2013).
34. Fritsche, L.G. *et al.* A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. *Nat Genet* **48**, 134-43 (2016).
35. Grassmann, F. et al. Genetic pleiotropy between age-related macular degeneration and 16 complex diseases and traits. *Genome Med* **9**, 29 (2017).
36. Browning, S.R. & Browning, B.L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet* **81**, 1084-97 (2007).
37. Choi Y., *et al.* Comparison of phasing strategies for whole human genomes. *PLoS Genet* **14**, e1007308 (2018).
38. Wang K., *et al.* ANNOVAR: Functional annotation of genetic variants from next-generation sequencing data. *Nucleic Acids Res* **38**, e164 (2010).
39. Ma C., *et al.* Recommended joint and meta-analysis strategies for case-control association testing of single low-count variants. *Genet Epidem* **37**, 539-550 (2013).
40. Zhan X., *et al.* RVTESTS: an efficient and comprehensive tool for rare variants association analysis using sequence data. *Bioinformatics* **32**, 1423-1426 (2016).
41. Wu M.C., *et al.* Rare-variant association testing for seqeuence data with the sequence kernel association test. *Am J Hum Genet* **89**, 82-93 (2011).
42. Sokal, R.R. & Rohlf, F.J. *Biometry: the principles and practice of statistics in biological research, 3rd Ed.* San Francisco: W.H. Freeman (1995).
43. Sanjak, J.S., Long, A.D. & Thornton, K.R. A Model of Compound Heterozygous, Loss-of-Function Alleles Is Broadly Consistent with Observations from Complex-Disease GWAS Datasets. *PLoS Genet* **13**, e1006573 (2017).
44. dos Reis, M. *et al.* Uncertainty in the Timing of Origin of Animals and the Limits of Precision in Molecular Timescales. *Curr Biol* **25**, 2939-50 (2015).
45. Riemer J., *et al.* Colorimetric ferrozine-based assay for the quantitation of iron in cultured cells. *Anal Biochem* **331**, 370-375 (2004).
46. Healy S., *et al.* Modelling iron mismanagment in neurodegenerative disease in vitro: paradigms, pitfalls, possibilities & practical considerations. *Progress in Neurobiol* **158**, 1-14 (2017).
47. Recalcati, S., Locati, M., Gammella, E., Invernizzi, P. & Cairo, G. Iron levels in polarized macrophages: regulation of immunity and autoimmunity. *Autoimmun Rev* **11**, 883-9 (2012).
48. Bowlus, C.L. The role of iron in T cell development and autoimmunity. *Autoimmun Rev* **2**, 73-8 (2003).
49. Dixon, S.J. *et al.* Ferroptosis: an iron-dependent form of nonapoptotic cell death. *Cell* **149**, 1060-72 (2012).
50. Torti, S.V. & Torti, F.M. Iron and cancer: more ore to be mined. *Nat Rev Cancer* **13**, 342-55 (2013).
51. Derynck, R., Akhurst, R.J. & Balmain, A. TGF-beta signaling in tumor suppression and cancer progression. *Nat Genet* **29**, 117-29 (2001).

52. Wakefield, L.M. & Roberts, A.B. TGF-beta signaling: positive and negative effects on tumorigenesis. *Curr Opin Genet Dev* **12**, 22-9 (2002).
53. Collinet, C. *et al.* Systems survey of endocytosis by multiparametric image analysis. *Nature* **464**, 243-9 (2010).
54. Kanai, M., *et al*. Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nat Genet* **50**, 390-400 (2018).
55. Pickrell, J.K., *et al*. Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet* **48**, 709-717 (2016).
56. Anttila, V., *et al*. Genome-wide meta-analysis identifies new susceptibility loci for migraine. *Nat Genet* **45**, 912-917 (2013).
57. Overbeek, M.J. *et al.* Pulmonary arterial hypertension in limited cutaneous systemic sclerosis: a distinctive vasculopathy. *Eur Respir J* **34**, 371-9 (2009).
58. Li, Y. *et al.* Heparin binding preference and structures in the fibroblast growth factor family parallel their evolutionary diversification. *Open Biol* **6**, 150275 (2016).
59. Itoh, N. & Ornitz, D.M. Fibroblast growth factors: from molecular evolution to roles in development, metabolism and disease. *J Biochem* **149**, 121-130 (2011).
60. Turner, N. & Grose, R. Fibroblast growth factor signaling: from development to cancer. *Nat Rev Cancer* **10**, 116-129 (2010).
61. De Domenico, I., *et al.* Hepcidin regulation: ironing out the details. *J Clin Invest* **117**, 1755-1758 (2007)
62. Bartnikas, T.B. & Fleming, M.D. A tincture of hepcidin cures all: the potential for hepcidin therapeutics. *J Clin Invest* **120**, 4187-4190 (2010).
63. Wrighting, D.M. & Andrews, N.C. Interleukin-6 induces hepcidin expression through STAT3. *Blood* **108**, 3204-3209 (2006).

Tables

Table 1: Significant genes identified by recessive diplotype scanning.

CHR	GENE	P	OR#	SNPs	Case+	Case-	Control+	Control-
chr6	*HFE*	1.29×10^{-8}	28.6	14	8	10	189	6707
chr12	*FGF6*	1.99×10^{-6}	22.8	10	6	12	153	6743
chr21	*KRTAP15-1*	7.55×10^{-5}	6.78	5	11	7	1271	5625
chr20	*XKR7*	1.18×10^{-4}	43.6	7	3	15	35	6861
chr20	*CABLES2*	1.28×10^{-4}	42.4	7	3	15	36	6860
chr22	*THOC5*	1.38×10^{-4}	6.24	9	13	5	1945	4951

The six most significant genes identified in the recessive diplotype scan are displayed. P-values are from a two-tailed Fisher's exact test. OR# shown as Haldane odds ratio. SNPs: number of genotyped variants per gene that were polymorphic in the samples studied. Case+: number of iron overload case individuals carrying recessive diplotypes with putative functional alleles. Case-: number of cases carrying at least one homolog at the gene without a putative functional allele. Control+: number of control individuals carrying recessive diplotypes with putative functional alleles. Control-: number of controls carrying at least one homolog at the gene without a putative functional allele.

Figure legends

Figure 1. Manhattan plot of the gene-based recessive diplotype association results. The association *P*-value testing hemochromatosis association for each gene ($-log_{10}p$ plotted on the ordinate) on different chromosomes is shown in alternating navy blue and yellow along the abscissa, with the experiment-wise significance level for the gene-based analyses across the exome (experiment-wise $\alpha = 3.14 \times 10^{-6}$) depicted in red.

Figure 2. Protein sequence alignment for FGF-4, FGF-5 and FGF-6 with heparin and FGFR binding domains. Protein domains summarized from a previous FGF-6 functional study[58]. Alignment and Heparin and FGFR-binding sites/regions (HBS and FGFR-BR, respectively) are shown for FGF-4, FGF-5 and FGF-6 proteins.

Figure 3. FGF-6 Active protein dosage effect on intracellular iron concentration. A ferrozine assay was applied for the evaluation of total cell iron content in HepG2 (Human liver hepatocellular carcinoma cell line), 786-O (Human kidney adenocarcinoma cell line), HCT-8 (Human ileocecal colorectal adenocarcinoma cell line), HCT116 (Human colon carcinoma cell line) and HFF-1 (Human skin fibroblast cell line) with 10μM FAC and 500μM ascorbate in cell culture media, respectively, with different concentrations of FGF-6 active protein (0ng/ml, 2.5ng/ml, 25ng/ml, and 250ng/ml). Control group was treated with ascorbate alone. After 48h incubation, cells were lysed and iron contents were determined with the ferrozine assay. **(A)** Total iron content in HepG2 cells with increasing FGF-6 protein concentration. **(B)** Total iron content in 786-O cells with increasing FGF-6 protein concentration. **(C)** Total iron content in HCT-8 cells with increasing FGF-6 protein concentration. **(D)** Total iron content in HCT-116 cells with increasing FGF-6 protein concentration.

Figure 4. The effect of *FGF6* nonsynonymous variants on hepcidin expression and intracellular iron concentration. (A) The effect of FGF-6 active protein treatment on mRNA expression of several iron metabolism genes in HepG2 liver hepatocellular carcinoma cell culture media compared to control. Protein concentration was 250ng/ml and the incubation time was 24h with 10μM FAC and 500μM ascorbate in the cell culture media. *HAMP* encodes for hepcidin. *HDAC2* encodes for histone deacetylase 2. *HMOX1* encodes for heme oxygenase 1. *TFRC* encodes for transferrin receptor 1 and HEPH encodes hephaestin. mRNA expression was quantified relative to *GAPDH* expression. Treatment with PBS served as control. A t-test was used test for pairwise differences between sets of observations. *P□0.05; **P□0.01. Results are the mean±SD of 3 observations in a single experiment. **(B)** Iron metabolism gene expression changes with *FGF6* mRNA transfection in the HepG2 cell culture media after 24h. Vector without *FGF6* served as control. A t-test was used to test for pairwise differences between sets of observations. *P□0.05; **P□0.01. Results are the

mean±SD of 3 observations in a single experiment. **(C, D)** Iron metabolism gene expression changes after the transfection by *FGF6* mRNA into various cell types with wildtype (WT) and the identified variants E172X (M1), D174V (M2) and R188Q (M3). Cell lines: HepG2 are liver hepatocellular carcinoma cells, HCT116 are ileocecal colorectal adenocarcinoma cells and HFF-1 are human normal skin fibroblasts. A t-test was used to test for pairwise differences between sets of observations. *P□0.05; **P□0.01. Results are the mean±SD of 3 observations in a single experiment. **(E, F)** Total intracellular iron concentration changes after the transfection with *FGF6* mRNA into three cell types with WT and the identified M1, M2, and M3 variants in the presence of FAC for 48h. A t-test was used to test for pairwise differences between sets of observations. *P□0.05; **P□0.01. Results are the mean±SD of 3 observations in a single experiment. **(G, H)** Ferritin protein level changes after the transfection by *FGF6* mRNA into the three cell types with WT and the identified M1, M2, and M3 variants in the presence of FAC for 48h. A t-test was used to test for pairwise differences between sets of observations. *P□0.05; **P□0.01. Results are the mean±SD of 3 observations in a single experiment.

Figure 5. Perls' stain and Ferritin expression. (A) FGF-6 protein level was evaluated by immunohistochemistry assay (IHC) in skin tissues from systemic sclerosis (SSc) patients and healthy controls (Normal). Staining was visualized by Nikon microscopy. A t-test was used to test for pairwise differences between average optical density (AOD) values between SSc and Normal observations. The ratio of positive stain areas to the total area was used to evaluate protein levels. AOD were quantified by Image J software. **P□0.01. (B) IHC with Perls' Prussian Blue stain for Ferritin protein was applied to evaluate the iron deposition in SSc skin tissues and Normal skin tissue. AOD values were quantified by Image J software. Staining was visualized by Nikon microscopy. **P□0.01. (C) IHC of FGF-6 protein in liver cancer tissue and control tissue. AOD were quantified by Image J Software. Staining was visualized by Nikon microscopy. *P□0.05. (D) IHC of Ferritin protein using Perls' Prussian Blue stain in liver cancer tissue and control tissue. AOD were quantified by Image J Software. Staining was visualized by Nikon microscopy. *P□0.05.

Figure 6. The proposed mechanism of FGF-6 in the regulation of hepcidin expression and iron concentrations. Paracrine FGF-6 interacts with FGFR with heparin or heparin sulphate proteoglycan (HPSG) as the cofactor to initial FGF pathway.[59] Activated FGFRs have the ability to phosphorylate specific tyrosine residues and activate STAT3 pathway.[60] Iron overload, and inflammation could positively regulate hepcidin by BMP/Smad pathway[61] and inflammatory IL-6/STAT3 pathways[62,63]. However, loss-of-function *FGF6* variants will silence the FGF6-FGFR pathway, increase free heparin and reduce expression of hepcidin, thereby decreasing the inhibition of ferroportin-mediated iron transfer from the intracellular compartment to the blood (i.e. increasing plasma levels of iron). In the SSc patients, IL-6 is increased so that hepcidin will be positively regulated which suppresses iron release to the plasma generating higher iron levels in skin cells.

Analysis Code:

We uploaded the Code to github: https://github.com/Shicheng-Guo/marshfield/blob/master/2ALOF/readme.md

Figure 1

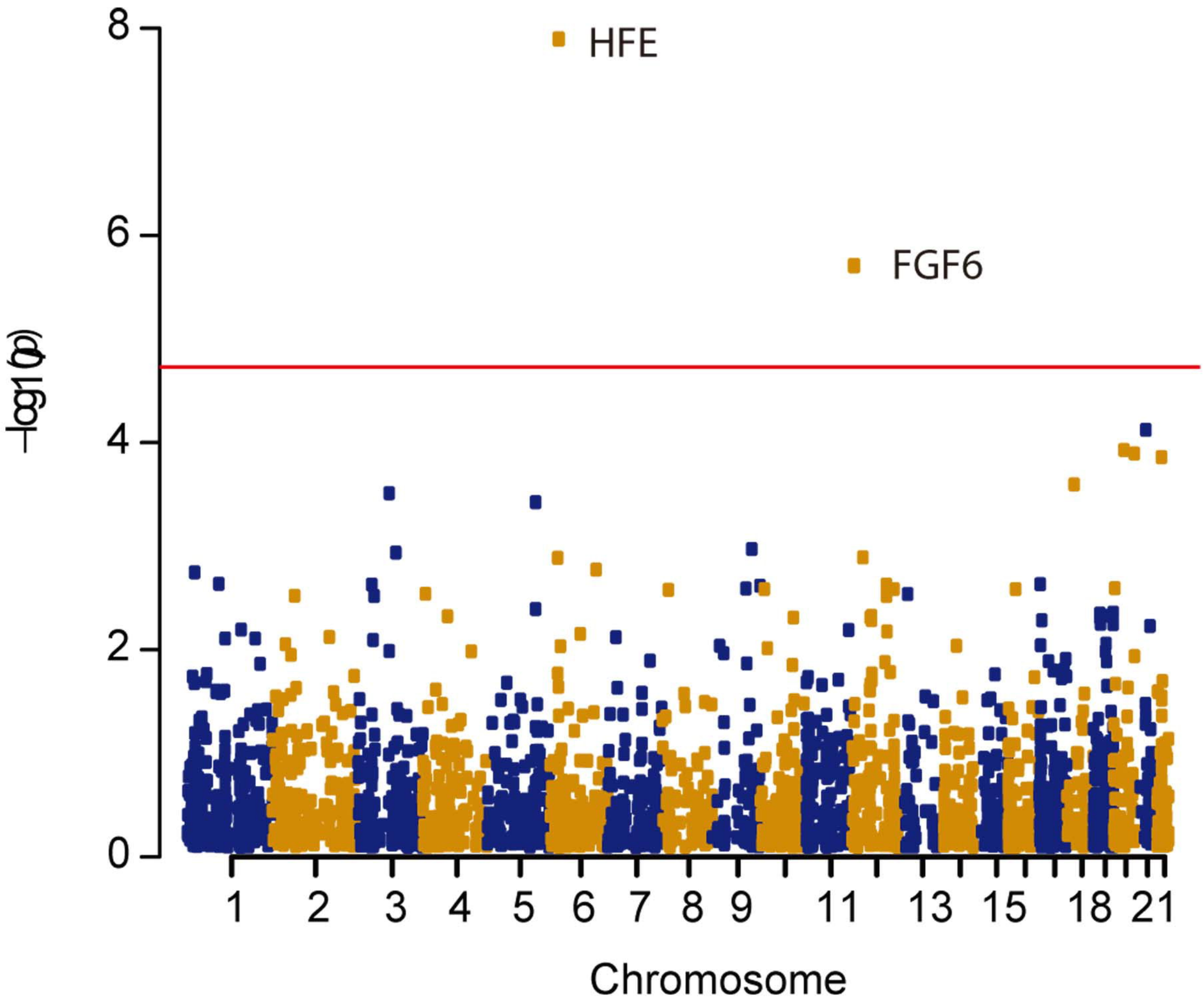


Figure 2

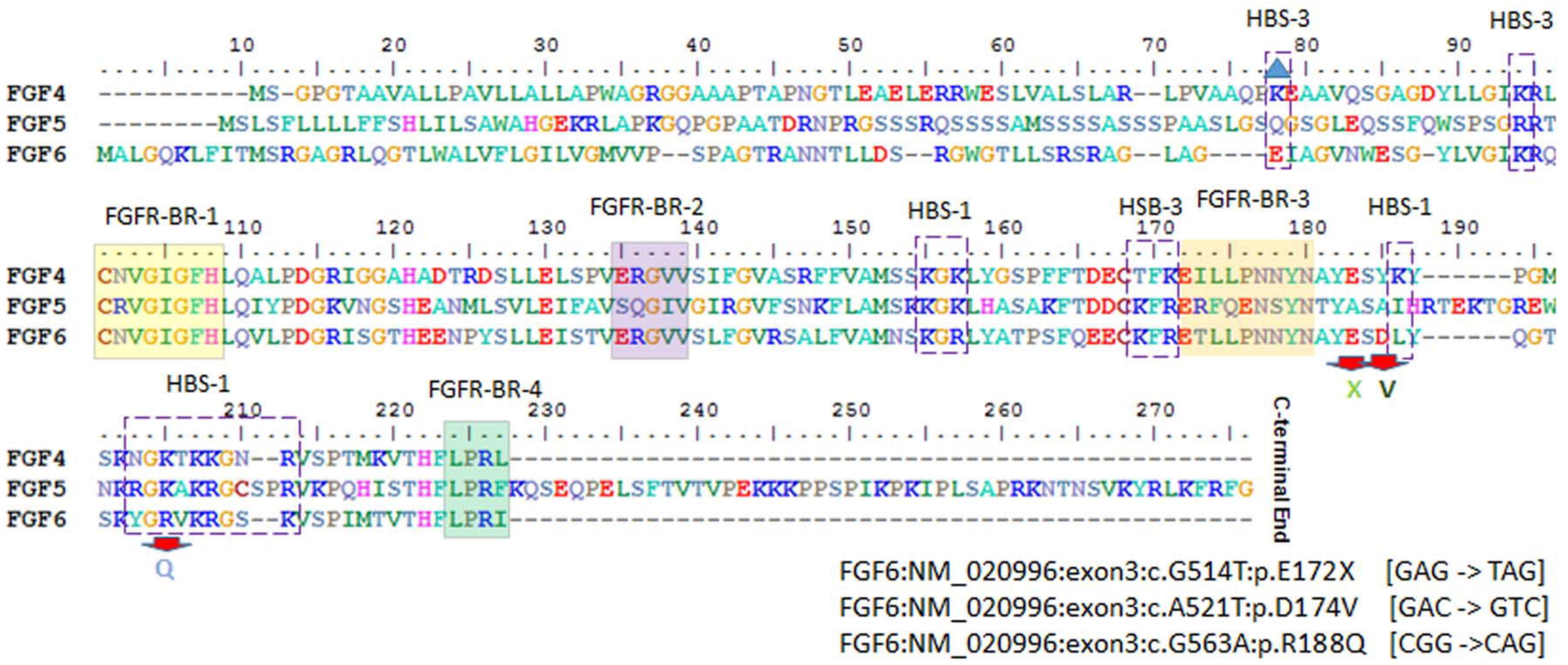
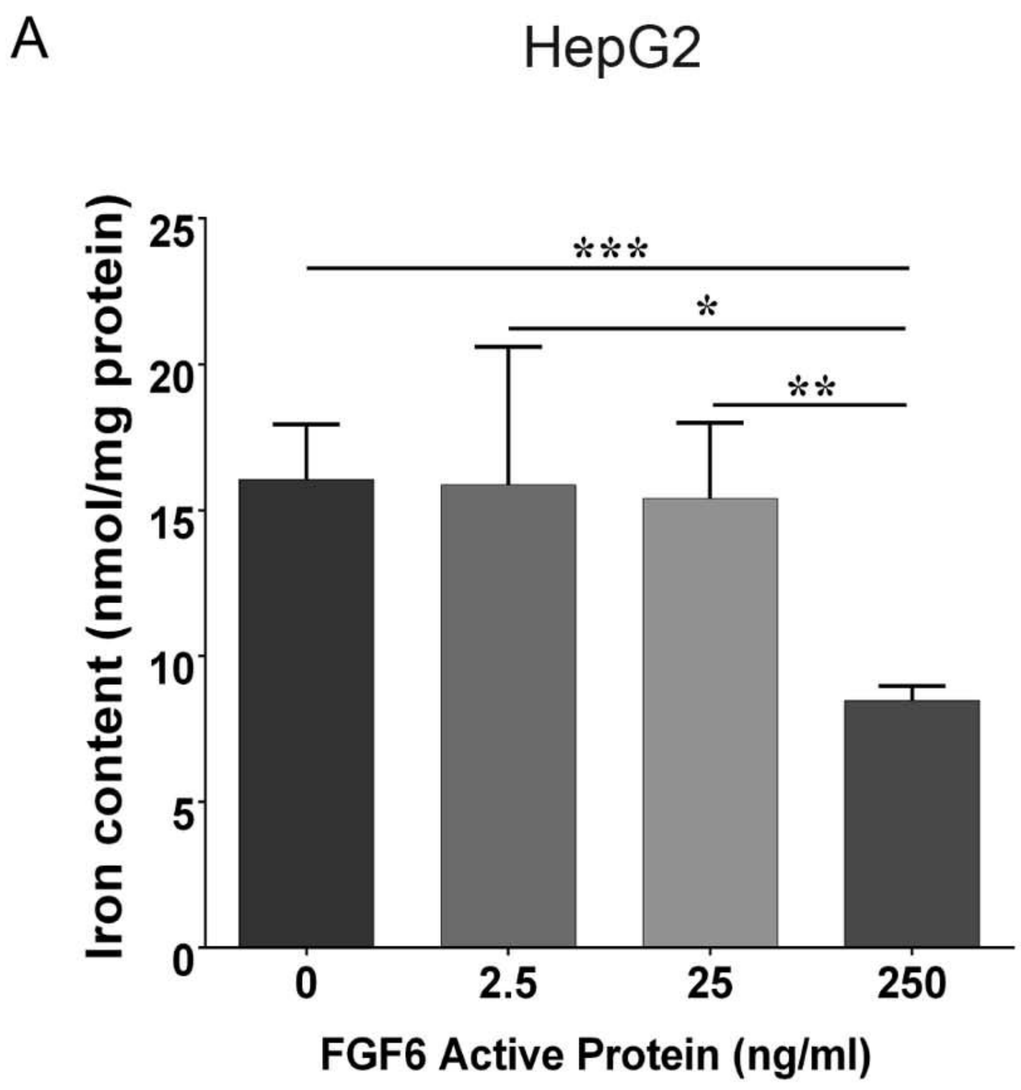
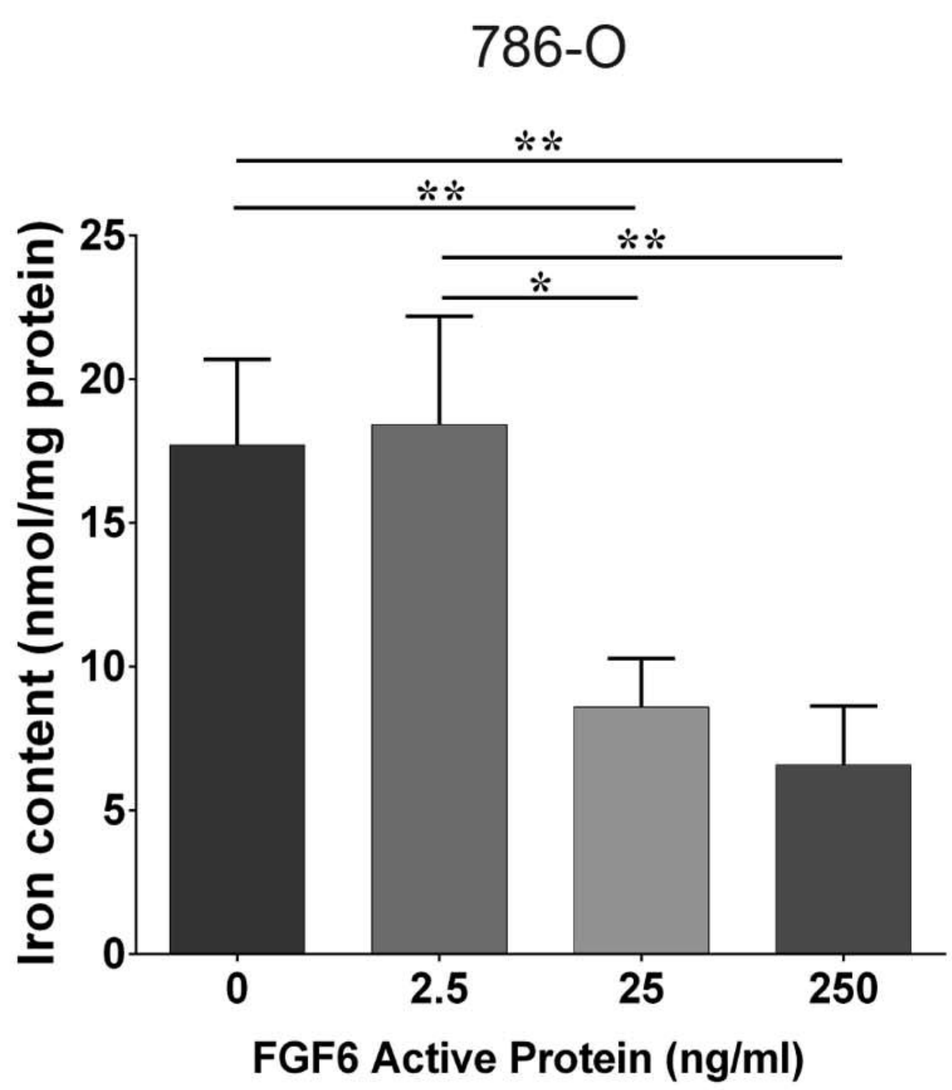
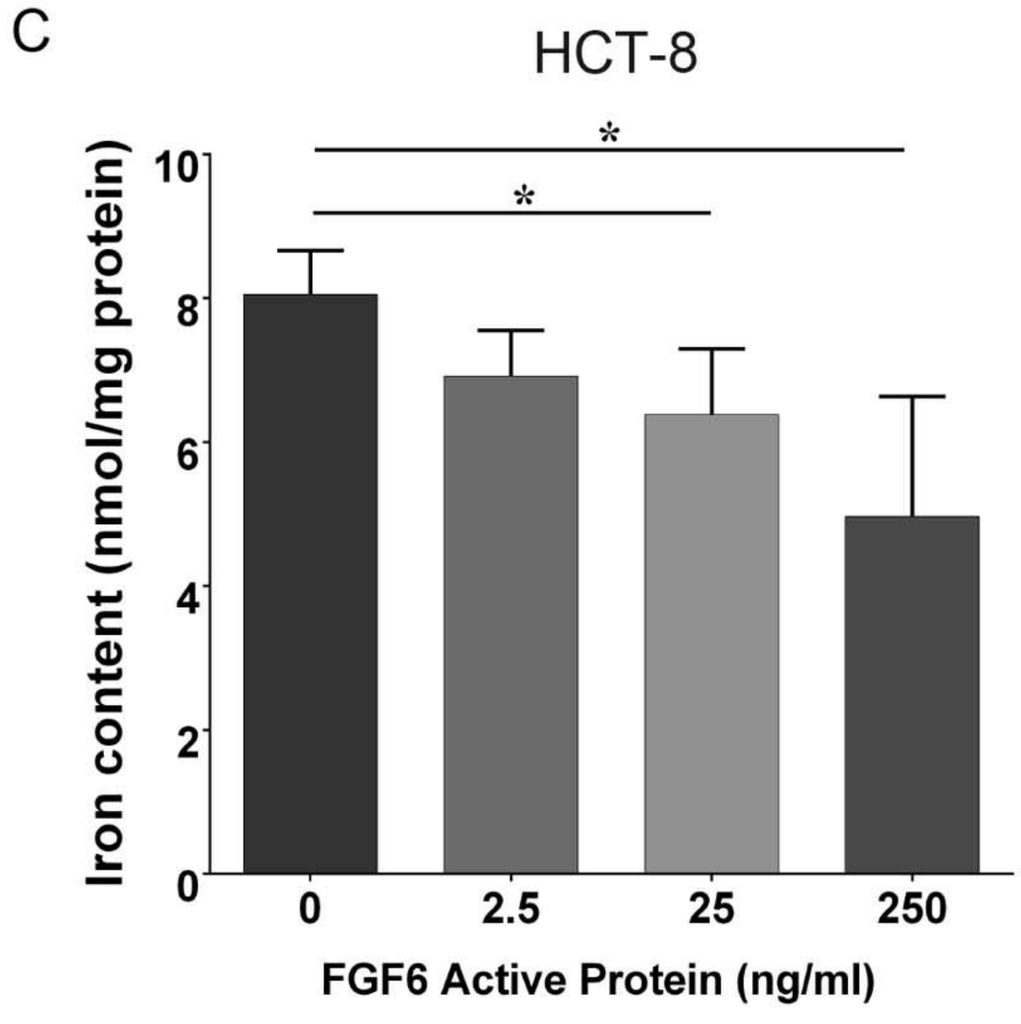

HBS-3
HBS-3
FGFR-BR-1
FGFR-BR-2
HBS-1
HSB-3
FGFR-BR-3
HBS-1
HBS-1
FGFR-BR-4
FGF4
FGF5
FGF6
X V
C-terminal End
Q
FGF6:NM_020996:exon3:c.G514T:p.E172X [GAG -> TAG]
FGF6:NM_020996:exon3:c.A521T:p.D174V [GAC -> GTC]
FGF6:NM_020996:exon3:c.G563A:p.R188Q [CGG ->CAG]

Figure 3


A
HepG2
Iron content (nmol/mg protein)
25
20
15
10
5
0
***
*
**
0
2.5
25
250
FGF6 Active Protein (ng/ml)


B
786-O
Iron content (nmol/mg protein)
25
20
15
10
5
0
**
**
**
*
0
2.5
25
250
FGF6 Active Protein (ng/ml)


C
HCT-8
Iron content (nmol/mg protein)
10
8
6
4
2
0
*
*
0
2.5
25
250
FGF6 Active Protein (ng/ml)

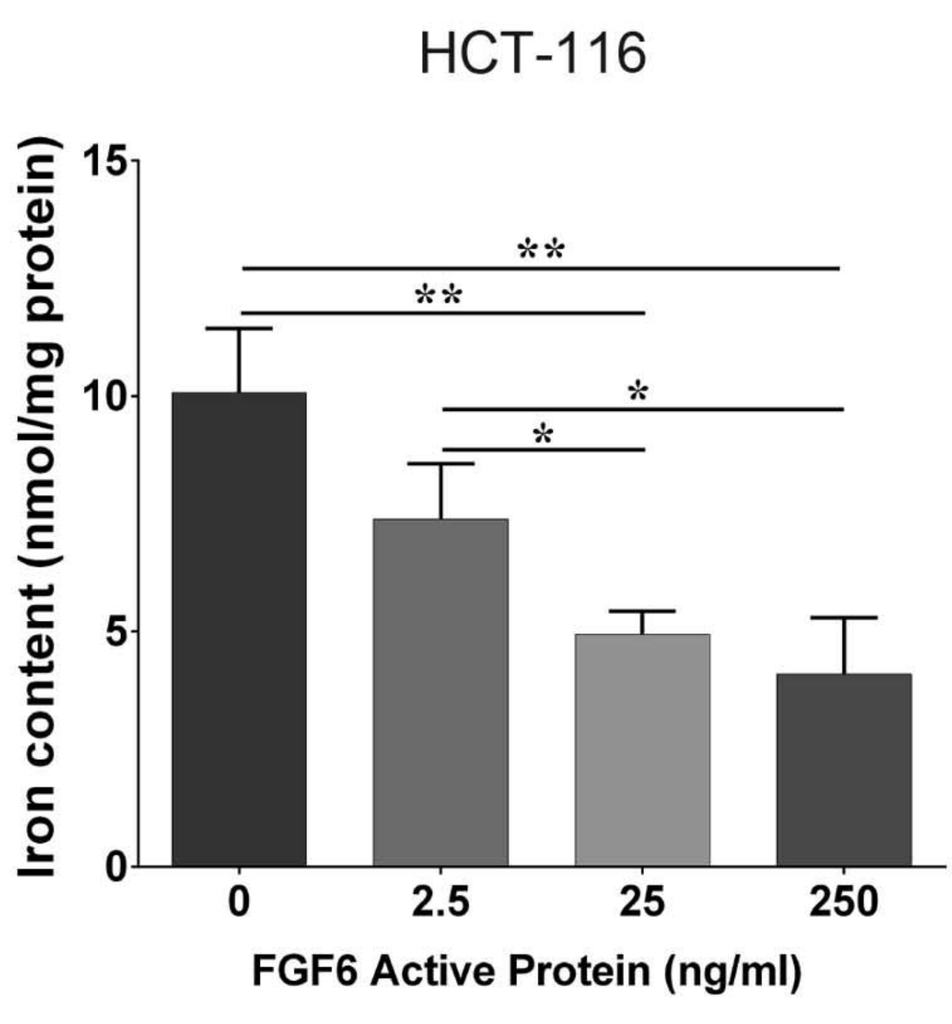

D
HCT-116
Iron content (nmol/mg protein)
15
10
5
0
**
**
**
*
*
0
2.5
25
250
FGF6 Active Protein (ng/ml)

Figure 4

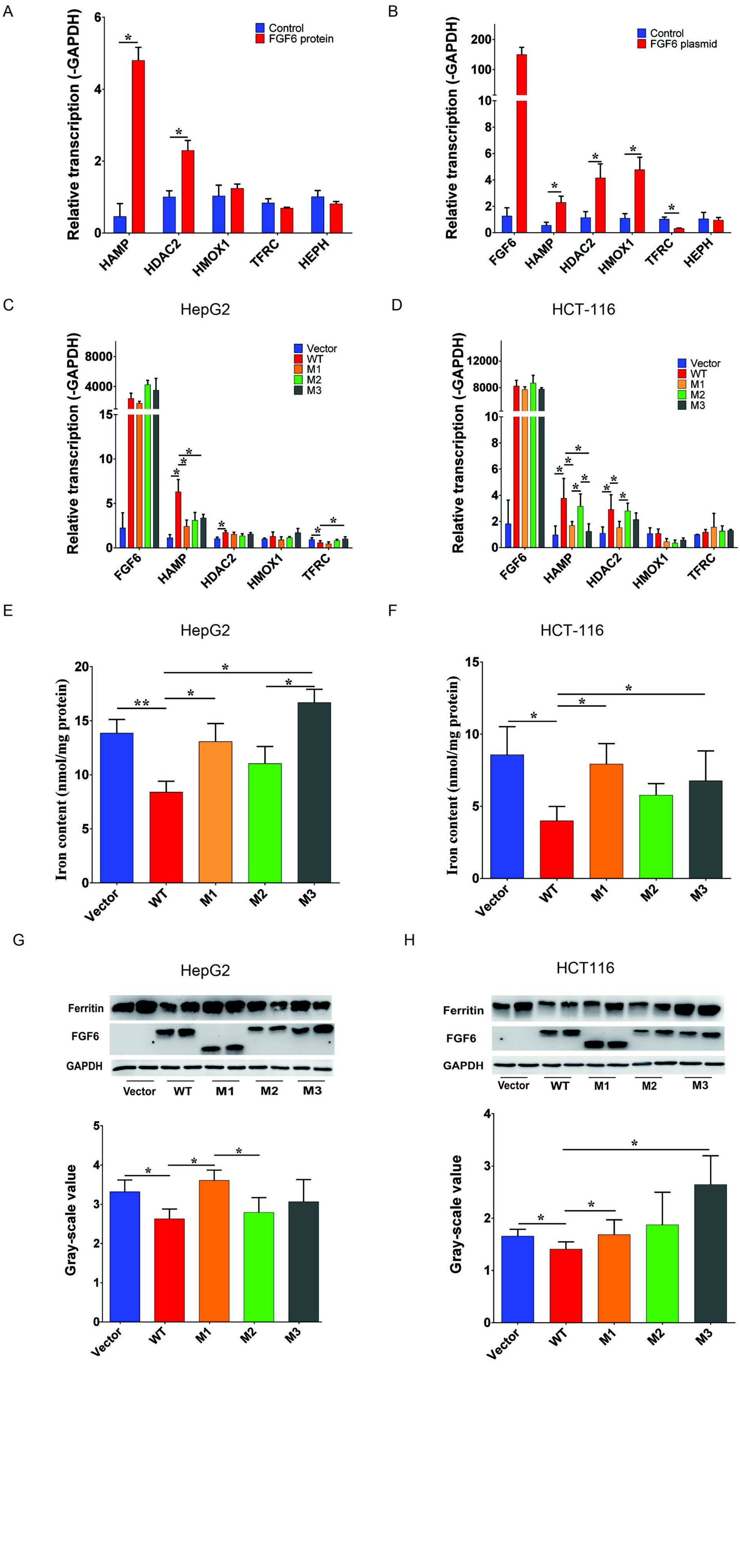

A
Relative transcription (-GAPDH)
Control
FGF6 protein
HAMP
HDAC2
HMOX1
TFRC
HEPH
B
Relative transcription (-GAPDH)
Control
FGF6 plasmid
FGF6
HAMP
HDAC2
HMOX1
TFRC
HEPH
C
HepG2
Relative transcription (-GAPDH)
Vector
WT
M1
M2
M3
FGF6
HAMP
HDAC2
HMOX1
TFRC
D
HCT-116
Relative transcription (-GAPDH)
Vector
WT
M1
M2
M3
FGF6
HAMP
HDAC2
HMOX1
TFRC
E
HepG2
Iron content (nmol/mg protein)
Vector
WT
M1
M2
M3
F
HCT-116
Iron content (nmol/mg protein)
Vector
WT
M1
M2
M3
G
HepG2
Ferritin
FGF6
GAPDH
Vector
WT
M1
M2
M3
Gray-scale value
Vector
WT
M1
M2
M3
H
HCT116
Ferritin
FGF6
GAPDH
Vector
WT
M1
M2
M3
Gray-scale value
Vector
WT
M1
M2
M3

Figure 5

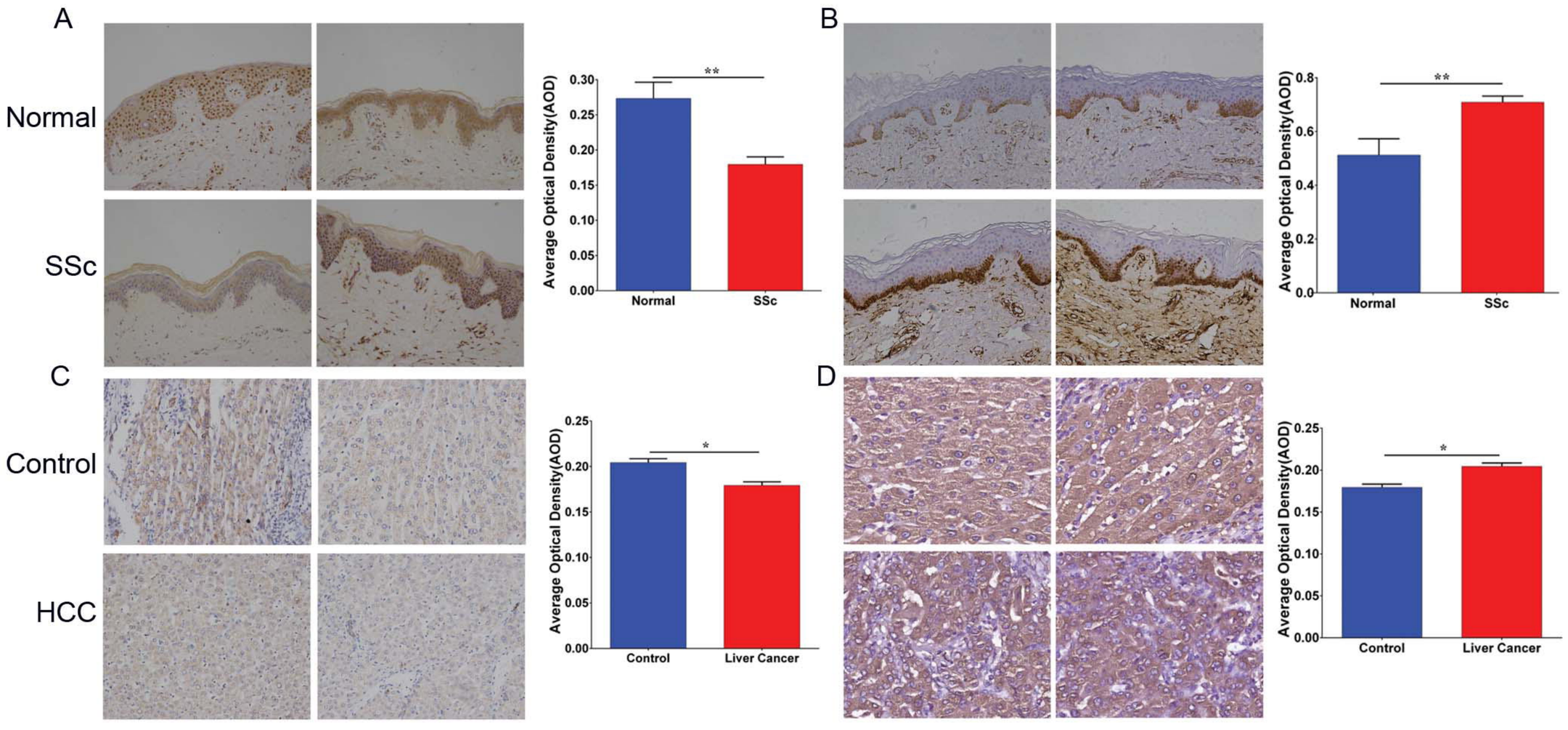
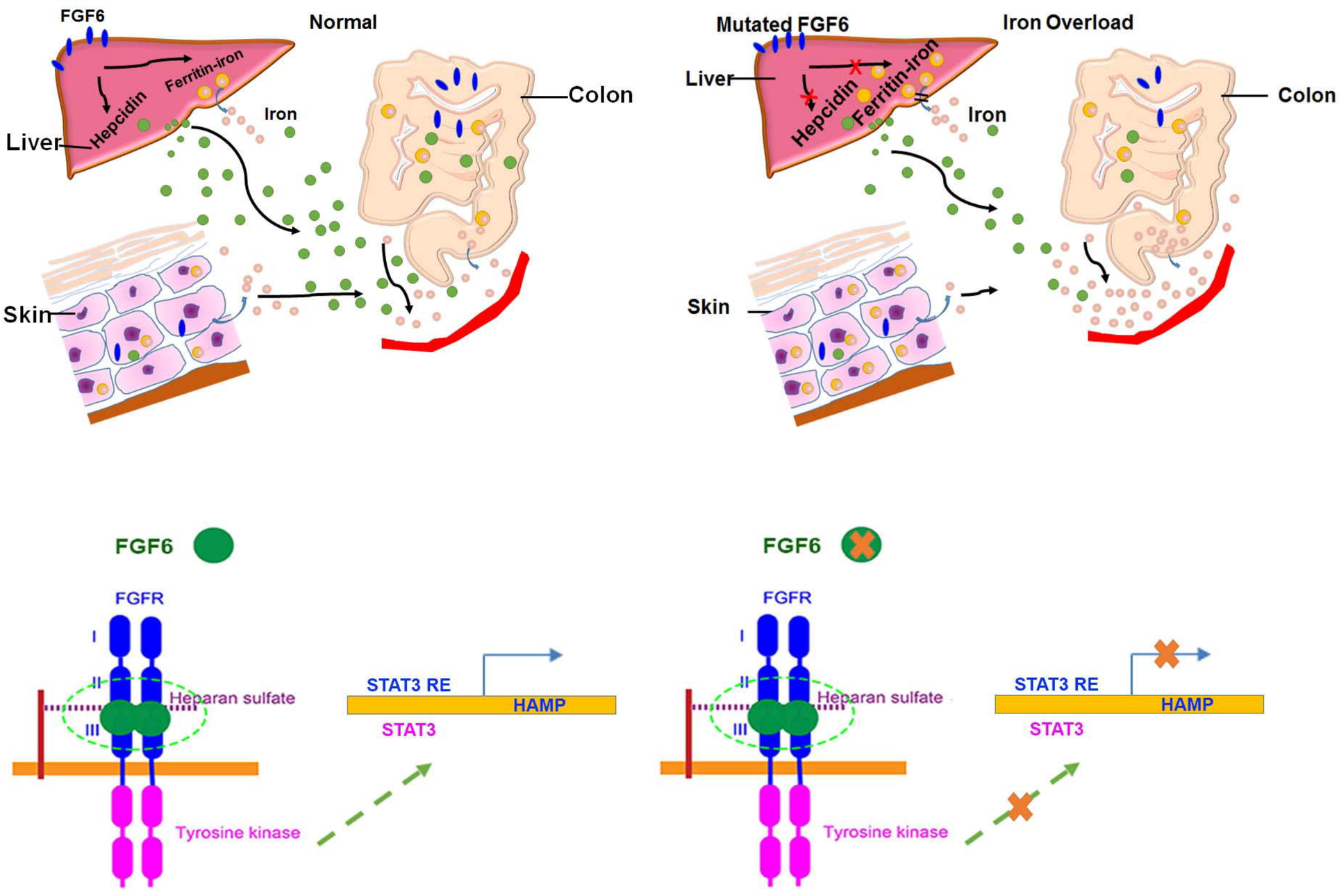

A
Normal
SSc
Average Optical Density(AOD)
0.30
0.25
0.20
0.15
0.10
0.05
0.00
**
Normal
SSc
B
Average Optical Density(AOD)
0.8
0.6
0.4
0.2
0.0
**
Normal
SSc
C
Control
HCC
Average Optical Density(AOD)
0.25
0.20
0.15
0.10
0.05
0.00
*
Control
Liver Cancer
D
Average Optical Density(AOD)
0.25
0.20
0.15
0.10
0.05
0.00
*
Control
Liver Cancer

Figure 6


FGF6
Normal
Ferritin-iron
Hepcidin
Iron
Liver
Colon
Skin
Mutated FGF6
Iron Overload
FGF6
FGFR
I
II
III
Heparan sulfate
Tyrosine kinase
STAT3 RE
HAMP
STAT3