## Supplementary Information

## 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 Guo1#, Shuai Jiang2#, Narendranath Epperla3, Yanyun Ma2, Mehdi Maadooliat1,4, Zhan Ye5, Brent Olson5, Minghua Wang6, Terrie Kitchner1, Jeffrey Joyce1, Robert Strenn5, Joseph J. Mazza7, Jennifer K. Meece8, Wenyu Wu9, Li Jin2, Judith A. Smith10, Jiucun Wang2,11,\*, Steven J. Schrodi1,12,\*

1Center for Human Genetics, Marshfield Clinic Research Institute, Marshfield, WI, USA

2State Key Laboratory of Genetic Engineering, Collaborative Innovation Center for Genetics and Development, School of Life Sciences, Fudan University, Shanghai, China

3Division of Hematology, The Ohio State University, Columbus, OH, USA

4Department of Mathematics, Statistics and Computer Science, Marquette University, Milwaukee, WI, USA

5Biomedical Informatics Research Center, Marshfield Clinic Research Institute, Marshfield, WI, USA

6Department of Biochemistry and Molecular Biology, Medical College, Soochow University, Suzhou, China

7Clinical Research Center, Marshfield Clinic Research Institute, Marshfield, WI, USA

8Integrated Research and Development Laboratory, Marshfield Clinic Research Institute, Marshfield, WI, USA

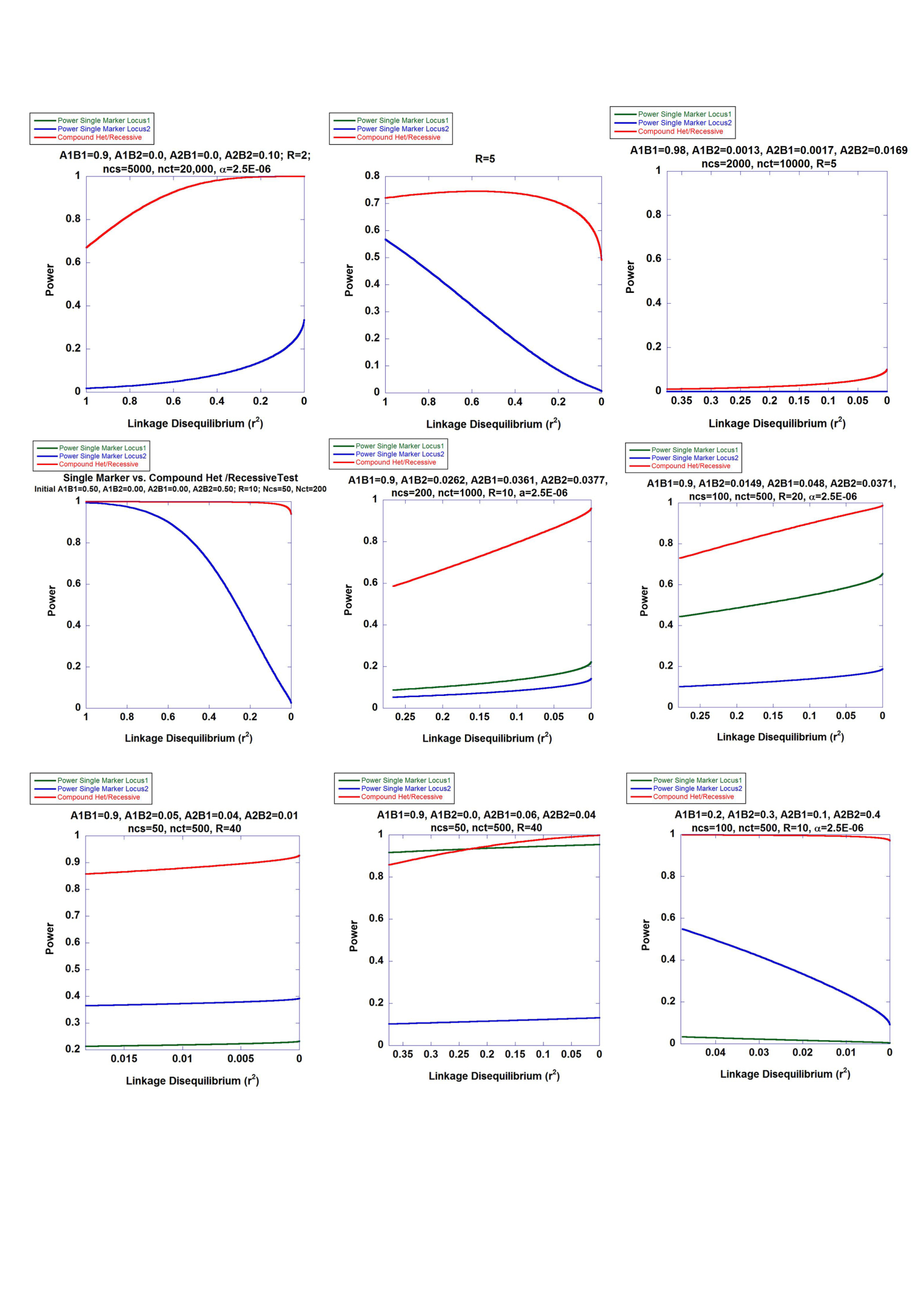
9Department of Dermatology, Huashan Hospital, Fudan University, Shanghai, China

10Department of Pediatrics, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI, USA

11Institute of Rheumatology, Immunology and Allergy, Fudan University, Shanghai, China

12Computation and Informatics in Biology and Medicine, University of Wisconsin-Madison, Madison, WI, USA

**Supplementary figure 1**



**Supplementary Figure 1. 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 explicity evaluating recessive diplotype effects. Baseline haplotype frequencies, case and control diploid sample sizes, and relative risk of disease-predisposing diplotypes parameters are shown. The of 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 1 error rate, adjusted for an exome-wide scan, was set to 2.5E-06 for all calculations.

**Supplementary figure 2**



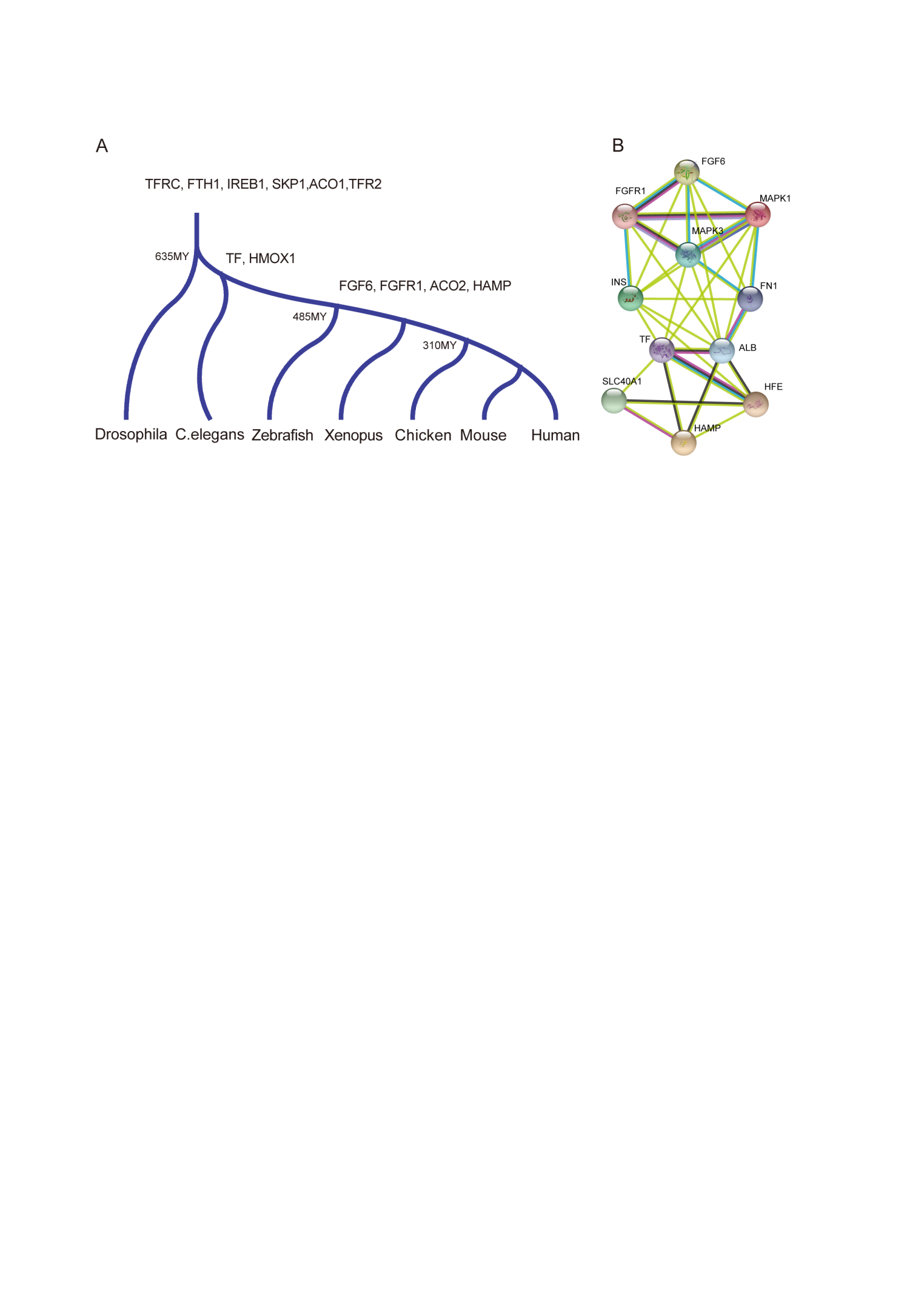
**Supplementary Figure 2. Minor Allele Frequency Distribution to PMRP dataset.** Distribution of minor allele frequency (MAF) to PMRP dataset, which was applied for gene-based recessive diplotype scanning in hemochromatosis analysis.

**Supplementary figure 3**



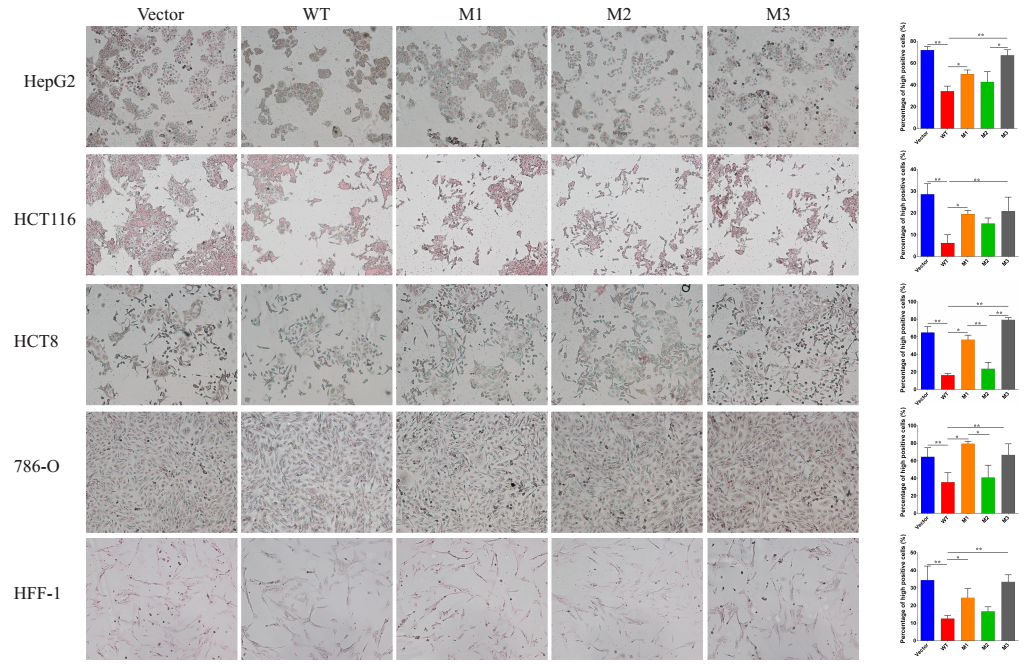
**Supplementary Figure 3.** **Quantile-Quantile Plot.** Q-Q plot for the exome-wide, gene-based recessive diplotype scanning in hemochromatosis is shown. Numerous genes had no recessive diplotypes with putative functional alleles and therefore yielded P-values of 1. The two data points exceeding the confidence interval represented *HFE* and *FGF6*.

**Supplementary figure 4**



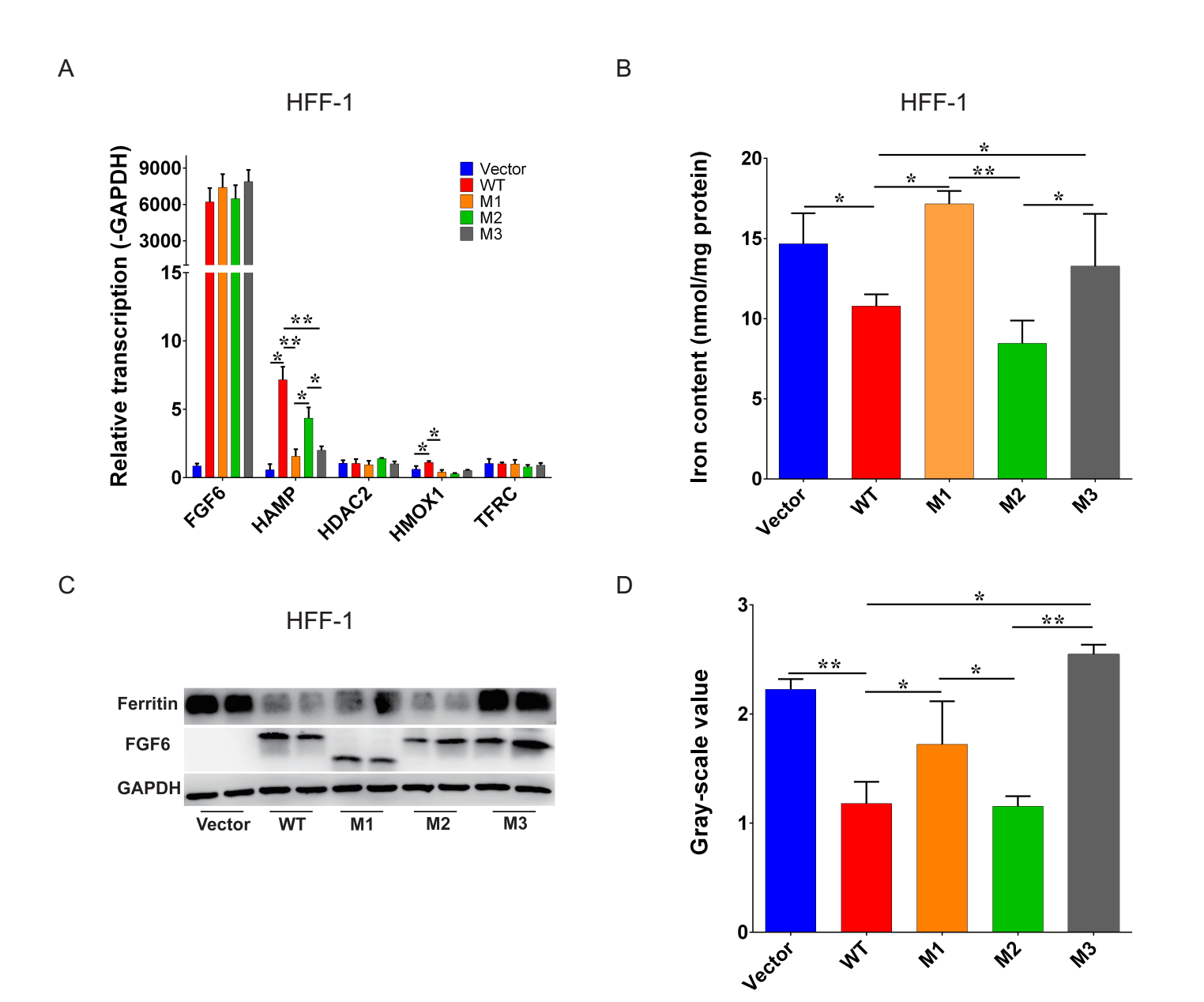
**Supplementary Figure 4. Comparative genomic analysis and protein-protein interaction (PPI).** The comparative genomic analyses revealed that *FGF6* evolved synchronously with other iron metabolism genes. **(A)** Main iron metabolism genes were collected and alignment was conducted to make the comparative 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).

**Supplementary figure 5**



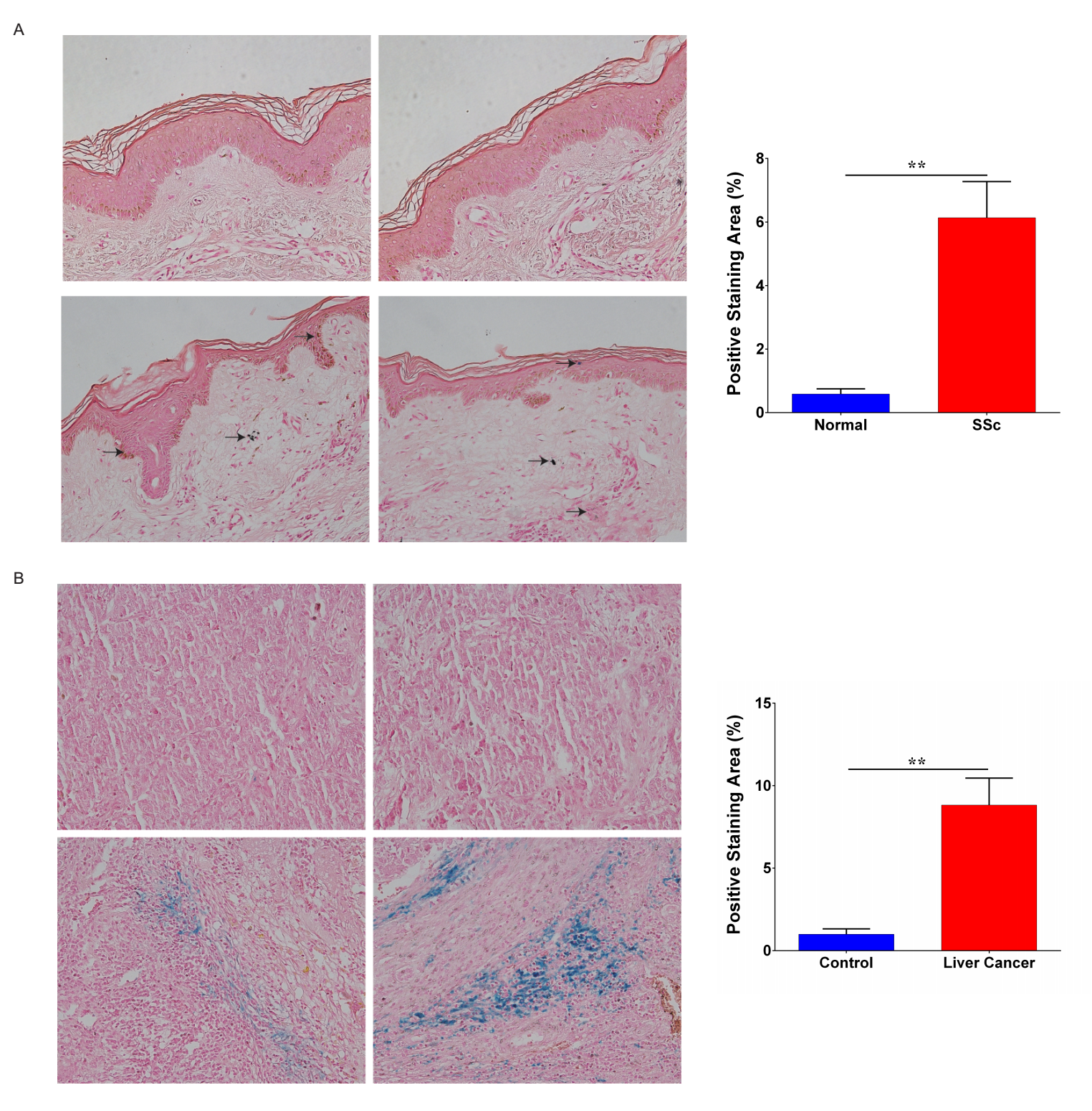
**Supplementary figure 5.** **Perls’ stain reveals that FGF6 loss-of-function nonsynonymous variants cause iron deposition.** Perls’ stain of various cell types (HepG2, HCT-116, HCT-8, 786-O and HFF-1) in the presence of FAC differs among transfection by FGF6 mRNA with wildtype and the identified variants R188Q, D174V and E172X.

**Supplementary figure 6**



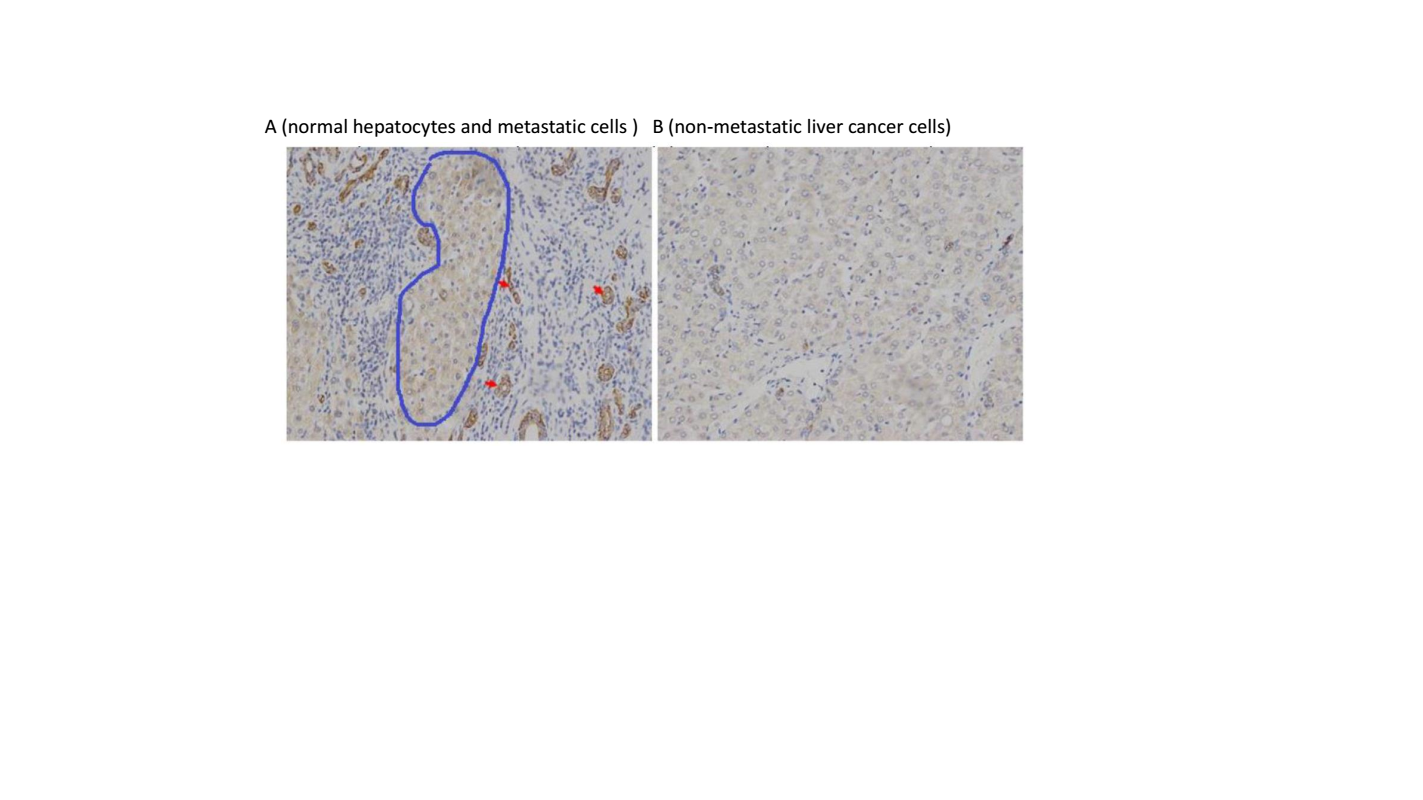
**Supplementary figure 6. FGF6 loss-of-function nonsynonymous variants cause hepcidin downregulation and iron deposition in HFF-1. (A)** Iron metabolism gene expression changes after the transfection by FGF6 mRNA into HFF-1 with wildtype and the identified variants R188Q, D174V and E172X. **(B)** Total 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 3 observations in 1 experiments.

**Supplementary figure 7**



**Supplementary figure 7. 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.

**Supplementary figure 8**



**Supplementary figure 8. FGF6 protein levels were different among normal, cancer and metastatic cells. (A)** IHC of FGF6 in normal hepatocytes and metastatic cells. The blue circle indicated normal liver tissue and the arrows indicated metastatic cells. (B) IHC of FGF6 in non-metastatic liver cancer cells.