

## Supplementary Methods

### **MTF1 binds to metal-responsive element e within the *ATP7B* promoter and is a strong candidate in regulating the *ATP7B* expression**

Amelie Stalke<sup>1,2</sup>, Eva-Doreen Pfister<sup>2</sup>, Ulrich Baumann<sup>2</sup>, Thomas Illig<sup>1,3</sup>, Eva Reischl<sup>4</sup>, Maria Sandbothe<sup>1</sup>, Beate Vajen<sup>1</sup>, Nicole Hüge<sup>1</sup>, Brigitte Schlegelberger<sup>1</sup>, Nils von Neuhoff<sup>1, #a</sup>, Britta Skawran<sup>1</sup>

<sup>1</sup>Department of Human Genetics, Hannover Medical School, Hannover, Germany

<sup>2</sup>Department of Pediatric Gastroenterology and Hepatology, Division of Kidney-, Liver and Metabolic Diseases, Hannover Medical School, Hannover, Germany

<sup>3</sup>Hannover Unified Biobank, Hannover Medical School, Hannover, Germany

<sup>4</sup>Research Unit of Molecular Epidemiology, Institute of Epidemiology II, Helmholtz Zentrum München, Neuherberg, Germany

<sup>#a</sup>Current Address: Clinic for Pediatrics III, Essen University Hospital, Essen, Germany

Corresponding author:

Amelie Stalke, email: Stalke.Amelie@mh-hannover.de

### **Cell culture**

Hepatocellular carcinoma cell line HLE was kindly provided by Professor Nam-Ho Huh (Department of Cell Biology, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, Okayama, Japan). The hepatocellular carcinoma cell line HepG2 cell line was purchased from ATCC (Manassas, VA, USA). Both cell lines were cultured in DMEM containing 10% FCS, 2 mM L-glutamine, and 100 U/mL penicillin/streptomycin at 37°C and 5% CO<sub>2</sub> in a humidified incubator.

### **Electrophoretic Mobility Shift Assay (EMSA)**

Nuclear extracts were prepared using NE-PER Nuclear and Cytoplasmic Extraction Reagents Kit (Thermo Fisher Scientific, Braunschweig, Germany). Protein concentration was determined by Bradford assay. 15 µg of nuclear extract or 50 ng of purified MTF1 (TP304861, Origene, Herford, Germany; purity >80 %, expressed in HEK293 cells, contains a C-terminal MYC/DDK tag) were incubated on ice for 25 min with 20 fmol double stranded HPLC-purified biotin-labeled oligonucleotides (Metabion, Planegg/Steinkirchen, Germany, see table S1) in a total volume of 20 µL in a buffer consisting of 10 mM Tris (pH 7.5), 100 mM KCl, 1 mM DTT, 0.8 µg poly [d(IC)], 0.05% NP-40, 5 mM MgCl<sub>2</sub>, 10 mM EDTA (not for MREe), 0.05 mM ZnSO<sub>4</sub> (only for MREe) and 2.5% Glycerol. For competition experiments nuclear extracts were preincubated with 4 pmol unlabeled HPLC-purified double stranded oligonucleotides (Metabion; see table S1) before adding biotin labeled probes. Samples were electrophoresed on ice on a 5% native polyacrylamide gel in 0.5x TBE buffer. DNA was transferred on a nylon membrane (Carl Roth, Karlsruhe, Germany) by tank blotting and crosslinked at 302 nm for 10 min. The membrane was developed with Chemiluminescent Nucleic Acid Detection Module Kit (Thermo Fisher Scientific).

**Table S1: Oligonucleotides for EMSA experiments**

Name	Sequence
MREc Bio	a 5' Biotin-AGT ACA GTG TCG <b>GAG CGC ACC</b> AGC GCG AGG T b 5' Biotin-ACC TCG CGC TGG <b>TGC GCT CCG</b> ACA CTG TAC T
MREc	a 5' AGT ACA GTG TCG <b>GAG CGC ACC</b> AGC GCG AGG T b 5' ACC TCG CGC TGG <b>TGC GCT CCG</b> ACA CTG TAC T
MREc mut I	a 5' AGT ACA GTG TCG <b>GAG CAT</b> TCC AGC GCG AGG T b 5' ACC TCG CGC TGG <b>AAT GCT CCG</b> ACA CTG TAC T
MREc mut II	a 5' AGT ACA GTG TCG <b>TGT GAT</b> TCC AGC GCG AGG T b 5' ACC TCG CGC TGG <b>AAT CAC ACG</b> ACA CTG TAC T
MREc mut III	a 5' AGT ACA GTG TCG <b>TGT GGC ACC</b> AGC GCG AGG T b 5' ACC TCG CGC TGG <b>TGC CAC ACG</b> ACA CTG TAC T
MREc mut III.1	a 5' AGT ACA GTG TCG <b>GGT GGC ACC</b> AGC GCG AGG T b 5' ACC TCG CGC TGG <b>TGC CAC CCG</b> ACA CTG TAC T
MREc mut III.2	a 5' AGT ACA GTG TCG <b>GAA GGC ACC</b> AGC GCG AGG T b 5' ACC TCG CGC TGG <b>TGC CAT CCG</b> ACA CTG TAC T
MREc mut III.3	a 5' AGT ACA GTG TCG <b>GAG GGC ACC</b> AGC GCG AGG T b 5' ACC TCG CGC TGG <b>TGC CCT CCG</b> ACA CTG TAC T
MREc mut III.4	a 5' AGT ACA GTG TCG <b>TAG CGC ACC</b> AGC GCG AGG T b 5' ACC TCG CGC TGG <b>TGC GCT ACG</b> ACA CTG TAC T
MREc mut IV	a 5' AGT ACA AAT GAT <b>GAG CGC ACC</b> AGC GCG AGG T b 5' ACC TCG CGC TGG <b>TGC GCT CAT</b> CAT TTG TAC T
MREc mut V	a 5' AGT ACA GTG TCG <b>GAG CGC ATA</b> CAT TCG AGG T b 5' ACC TCG AAT GTA <b>TGC GCT CCG</b> ACA CTG TAC T
MREc mut V.1	a 5' AGT ACA GTG TCG <b>GAG CGC ATA</b> CAT GCG AGG T b 5' ACC TCG CAT GTA <b>TGC GCT CCG</b> ACA CTG TAC T
MREc mut V.2	a 5' AGT ACA GTG TCG <b>GAG CGC ATA</b> CAC GCG AGG T b 5' ACC TCG CGT GTA <b>TGC GCT CCG</b> ACA CTG TAC T
MREc mut V.3	a 5' AGT ACA GTG TCG <b>GAG CGC ATA</b> CGC GCG AGG T b 5' ACC TCG CGC GTA <b>TGC GCT CCG</b> ACA CTG TAC T
MREc mut V.4	a 5' AGT ACA GTG TCG <b>GAG CGC ATA</b> AGC GCG AGG T b 5' ACC TCG CGC TTA <b>TGC GCT CCG</b> ACA CTG TAC T
MREc mut V.5	a 5' AGT ACA GTG TCG <b>GAG CGC ATC</b> AGC GCG AGG T b 5' ACC TCG CGC TGA <b>TGC GCT CCG</b> ACA CTG TAC T
MREc mut V.6	a 5' AGT ACA GTG TCG <b>GAG CGC ACA</b> CAT TCG AGG T b 5' ACC TCG AAT GTG <b>TGC GCT CCG</b> ACA CTG TAC T
MREc mut V.7	a 5' AGT ACA GTG TCG <b>GAG CGC ACC</b> CAT TCG AGG T b 5' ACC TCG AAT GGG <b>TGC GCT CCG</b> ACA CTG TAC T
MREc mut V.8	a 5' AGT ACA GTG TCG <b>GAG CGC ACC</b> AAT TCG AGG T b 5' ACC TCG AAT TGG <b>TGC GCT CCG</b> ACA CTG TAC T
MREc mut V.9	a 5' AGT ACA GTG TCG <b>GAG CGC ACC</b> AGT TCG AGG T b 5' ACC TCG AAC TGG <b>TGC GCT CCG</b> ACA CTG TAC T
MREc mut V.10	a 5' AGT ACA GTG TCG <b>GAG CGC ACC</b> AGC TCG AGG T b 5' ACC TCG AGC TGG <b>TGC GCT CCG</b> ACA CTG TAC T
MREc mut VI	a 5' AGT ACA GTG TCG <b>TGT GGC ATA</b> CAT TCG AGG T b 5' ACC TCG AAT GTA <b>TGC CAC ACG</b> ACA CTG TAC
MREd Bio	a 5' Biotin-GGG AGG ACG GCG <b>GCG CGC AAC</b> TTT GAA TCA T b 5' ATG ATT CAA AGT <b>TGC GCG CCG</b> CCG TCC TCC C
MREd	a 5' GGG AGG ACG GCG <b>GCG CGC AAC</b> TTT GAA TCA T b 5' ATG ATT CAA AGT <b>TGC GCG CCG</b> CCG TCC TCC C
MREd mut II	a 5' GGG AGG ACG GCG <b>ATA TAT GAC</b> TTT GAA TCA T

	b 5' ATG ATT CAA AGT <b>CAT ATA</b> TCG CCG TCC TCC C
MREd mut III	a 5' GGG AGG ACG GCG <b>ATATGC</b> AAC TTT GAA TCA T b 5' ATG ATT CAA AGT <b>TGC ATATCG</b> CCG TCC TCC C
MREd mut IV	a 5' GGG AGG GTA ATA <b>GCG CGC</b> AAC TTT GAA TCA T b 5' ATG ATT CAA AGT <b>TGC GCG</b> CTA TTA CCC TCC C
MREd mut V	a 5' GGG AGG ACG GCG <b>GCG CGC</b> AGT CCC AAA TCA T b 5' ATG ATT <b>TGG GAC</b> <b>TGC GCG</b> CCG CCG TCC TCC C
MREe Bio	5' Biotin-GGC CTT <b>TGC GCA</b> CAG CGG ATC GAT TTT 5' AAA ATC GAT CCG CTG <b>TGC GCA</b> AAG GCC
MREe	5' GGC CTT <b>TGC GCA</b> CAG CGG ATC GAT TTT 5' AAA ATC GAT CCG CTG <b>TGC GCA</b> AAG GCC
MREe mut I	a 5' GGC CTT <b>TGC GTG</b> TAG CGG ATC GAT TTT b 5' AAA ATC GAT CCG CTA <b>CAC GCA</b> AAG GCC
MREe mut II	a 5' GGC CTT <b>CAT ATG</b> TAG CGG ATC GAT TTT b 5' AAA ATC GAT CCG CTA <b>CAT ATG</b> AAG GCC
MREe mut III	a 5' GGC CTT <b>CAT ACA</b> CAG CGG ATC GAT TTT b 5' AAA ATC GAT CCG CTG <b>TGT ATG</b> AAG GCC
MREe mut III.1	a 5' GGC CTT <b>TAT ACA</b> CAG CGG ATC GAT TTT b 5' AAA ATC GAT CCG CTG <b>TGT ATA</b> AAG GCC
MREe mut III.2	a 5' GGC CTT <b>TGT ACA</b> CAG CGG ATC GAT TTT b 5' AAA ATC GAT CCG CTG <b>TGT ACA</b> AAG GCC
MREe mut III.3	a 5' GGC CTT <b>TGC ACA</b> CAG CGG ATC GAT TTT b 5' AAA ATC GAT CCG CTG <b>TGT GCA</b> AAG GCC
MREe mut III.4	a 5' GGC CTT <b>CGC GCA</b> CAG CGG ATC GAT TTT b 5' AAA ATC GAT CCG CTG <b>TGC GCG</b> AAG GCC
MREe mut IV	a 5' AAT TCC <b>TGC GCA</b> CAG CGG ATC GAT TTT b 5' AAA ATC GAT CCG CTG <b>TGC GCA</b> GGA ATT
MREe mut V	a 5' GGC CTT <b>TGC GCA</b> CGA TAA GTC GAT TTT b 5' AAA ATC GAC TTA TCG <b>TGC GCA</b> AAG GCC
MRE-s	a 5' AGCTCGAGGGAGCTCTGCACACGCCCGAAAAGTGTCGA b 5' TCGACACTTTTCGGGCCGTGTGCAGAGCTCCCTCGAGCT

MRE-consensus sequence is shown in bold, mutated bases are highlighted in grey.

### Luciferase reporter assay with cotransfection of transcription factor candidates

To investigate if the transcription factors identified in MatInspector analysis are able to interact with the *ATP7B* promoter region, Luciferase reporter assays under coexpression of transcription factor candidates were performed. For this, a 2.5 kb fragment of the *ATP7B* promoter region was PCR-amplified from genomic DNA (see table S2), digested with *KpnI* and *NheI*, and cloned upstream of firefly luciferase into the pGL3-Basic vector (Promega, Mannheim, Germany). MREc, MREd or MREe were mutated using the QuikChange Lightning Site-Directed Mutagenesis kit (Agilent, Technologies, Santa Clara, CA, USA). Introduced mutations are highlighted in grey:

MREc: GTACAGTGTCTG**TGGCGCATAC**ACGCGAGGTGGCCGAGAC

MREd: GCCCCAGGTCGGGAGGACGGCG**ATATATG**ACTTTGAATCATCCGTGTGAAG

MREe: GTTGGAGGCCATTGGCTGGCCT**CATATG**TAGCGGATCGATTTTCCAGGTGCC

Coding sequences of potential transcription factors were obtained by expression vectors CMV-HA-hTAF6 (gift from Naoko Tanese, Addgene plasmid # 1798), pCMVHA E2F2 (gift from Kristian Helin, Addgene plasmid # 24226), pGEM-MTF1 (SinoBiological Inc., Peking, China), pGEM-TAF9

(SinoBiological Inc.) and RSV-Sp1 (gift from Robert Tjian, Addgene plasmid # 12098). Coding sequences were cloned into pcDNA3 vector (Invitrogen, Carlsbad, USA) using existing restriction sites if compatible (pCMVHA E2F2 with *Bam*H I; RSV-SP1 with *Eco*R I und *Xho* I) or were PCR-amplified from expression vectors by primers with overhangs to generate compatible restriction sites (pGEM-MTF1, CMV-HA-hTAF6, pGEM-TAF9) (see table S2).

For Lipofectamine 2000 (Life Technologies) transfection 40 ng pGL3 reporter construct, 3 ng pGL4.70-EF1 $\alpha$  and 40 ng for each pcDNA3 vector were filled up with MIGR1 (gift of Warren Pear, Addgene plasmid # 27490) to a total DNA amount of 130 ng and added to simultaneously seeded 24,000 HepG2 cells (white 96 well plates, 24,000 cells per well). Cotransfected pGL4.70 (Promega) with an EF1 $\alpha$  promoter inserted at the *Xho*I restriction site upstream of Renilla luciferase served for normalization. Luciferase activities were measured with the DualGlo Luciferase Assay System (Promega) 48 h after transfection. Normalized pGL3 Basic vector luciferase activity was set to 0, normalized pGL3 Promoter vector (Promega) activity was set to 1. All used vector inserts were verified by Sanger sequencing.

**Table S2: Primer sequences used to generate plasmid inserts**

Cloning	Primer	Sequence 5' to 3'	Restriction site
ATP7B-Promoter into pGL3	ATP7B K3.2 Fo	<u>TCCGGTACC</u> <u>GGATCCATTGAAGGGCCTTG</u>	<i>Kpn</i> I
	ATP7B K3.2 Re	<u>GCAGCTAGCC</u> <u>CGTCCCGCACGGACACC</u>	<i>Nhe</i> I
MTF1 into pcDNA3	MTF-1-ORF-HindIII Fo	<u>AACGTAAAGCTT</u> ATGGGGGAACACAGTCCAG	<i>Hind</i> III
	MTF-1-ORF-NotI Re	GTGTAC <u>GCGGCCGCT</u> CACTTGGAGAAGCTGCTGG	<i>Not</i> I
TAF6 into pcDNA3	TAF6-ORF-KpnI Fo	<u>ATTCTAGGTACC</u> ATGGCTGAGGAGAAGAAG CTG	<i>Kpn</i> I
	TAF6-ORF- XhoI Re	<u>GAGAACCTCGAGT</u> CACGGAGCAGGCTGA	<i>Xho</i> I
TAF9 into pcDNA3	TAF9-ORF-BamHI Fo	<u>AACGTAGGATCC</u> ATGGAGTCTGGCAAGACG	<i>Bam</i> H I
	TAF9-ORF-NotI-Re	<u>CTGTACGCGGCCGCT</u> TACAGATTATCATAGTCATCA TCAT	<i>Not</i> I

overhangs are highlighted in grey, restriction sites are underlined. (manufacturer: Metabion)

Upstream region of *ATP7B*, NC\_000013.10(NM\_000053.3), *Homo sapiens*, chr.13, build GRCh37/hg19

**GGATCCATTGAAGGGCCTTGG**

GACCTGCATGGTCCATGGAGCACAATTTGAGAACCACTGCCTTAAAGGATG  
AACAGAGTCTGACAAAAGTGGTAAGATTTATCATTTATTTAACATTTGGAATATATGAAT  
TCATAAAAGCAACACTACAGAGGACAGTTTGCTTTCTCAGGACACTCTTTCAAAGCTTA  
TTTAAATGCAGCTGACCCATATAATGTAAAGGTTTTCTGGCAATTCCTTCAAATAATT  
TAGGTGGGAAAATTCATGTGACAAGCTTTTCAAAGTAACTATTAACATTTAACTTAAA  
AGCTGCATAATAACTAAAGTTAAACCTGTCTTTAAGTCTGGTTAGTGAGTTTATTTAAA  
CATTATAATGTAAAGCACTGAGGAATATTACGTATGACAGGGAAGAACCAAAAGGCTCTG  
ATAACAGACTTTCTCATTCAAATGTTTAAAGTAATGATTACAGAATTTGAAGTAGAGTG  
GACAAAAACGAAGAAGAAATGCCTGAACCTTTTGAAGAAATTAGGTATATACTGCTATTCA  
GGAAGCTATGGTGTTAATTATAAATTTTACAGCTGAACAAAAATATGTAGGTGGTAAAAA  
AACTAGGTAGGTAGATATACTAGGCTTCTGTTTTATTAGTGGAGGGCAAGGCATTTAA  
AATACAGTAGTACTACAACAACAAAGGTGGAAAAAGCCTGGGAGTTTTTTCACCAAAAT  
GTTAACAGTGGCTATCTCCGGATGGTAGCATTCTGGGGTTTTTCTTTTACTTATCTA  
TTTTCTAATTTTTCTAGAAGAAACAGCATATACACGTAATACAAATAAAATAATTTTTA  
TGTAACCTTCTACATCATAAAAAATTAAGTAACAGATTATCAAAAGCAGAGGATTCAAC  
ATAAGCTTTATATTTAAGTGACGTGTTAACAATGGCAGTTTTGGTGAGTTTAAAGTAGA  
AACTAAGATGAGAAGGAAGATCCGGTTTTCTTTACCGAAAAGAGACATGATAGATTGG  
AAAATGTCTCGTGGCTCAATCGTAAAGAAAAATAAGGAATGAATATGAATAATTTATGAC  
AGACTAACTGGAGTGTGAAAGGGAAAAAAAAGCTGAAGCAGGGAAGGGAAGCTCGACCCC  
TCGGCCAAACAGTGAAGCAGAAGAAAAACGCGGTAAATTGAGTCTTTAAGACCTAGTCTATT  
CCAGTCTATATAGAAGACCCCCAAGAAAAAGAAAGACTGCTCATTGGCCTGAGAAATTAC  
CTTAGATTATGACAAGGAAGGCCATTTGCCCGCAAAATTTAGCTACACTGGACGGGCAA  
GTACCCTACAGAAGAGAAAACGTCTGTGAGCCACACGACCGGCTGCTCACCTCAACAA  
CTTGACAGGCACCAGCTCCTTTCCGCGCGCCCATCTTCCGCGACCCCCGAAGTCAAG  
AAACGCTTCACTTTCTTTTCCCTATTGGCTCCTGAGAAAGCAAGCCGTGCTCGCCCCGC  
CCCCACGGGCCAATTGTGCGTTACTATTGGTTACTGGTAGCCGCTTCCACGGCCTTCCA  
GCCAATAGAATATGCCGAGGCGTAGACTAGTGTTCCGCGTGGCGCACACGGCTCCCGCCC  
CCGTGGGCGGGACAGCAGTGGGGGTTGGCTGAGGAGGGCGTGGCCTGTGATTGACAGC  
CGTCGCTCCCTCCCTCGGCCACCTCCCCACTAGAAGCCCCCGCTGGGCGCCTGCGCCC  
CCGTTCCCGGCCCAAGCCCCGCGCCGTTGGAGGCCATTGGCTGGCCTT**TGCGCACAGC**  
GGATCGATTTTCCAGT**GCGG**AGTTCACTCTTGCCGCGTTGCTTCTTTGGGACCCACG  
GCGTCCGGCAGCCAGGCGCAGAGTCCGAGGAGGGGGCAGCGCAGAGCGGACCCGACGCGG  
CGCCGCCGGGCACCTTCCCCGACGGCGTGGGTGAGCCCTGGGAGCTGAGTCTGCGGGCC  
GGCTCTGCGCAGCTCACCTGCCCTCCCGCTCCCGCACACGCGTGAGATCCCAGTACAGTG  
TCGGAGCGCACACAGCGGAGGTGGCCGAGACCGCGGAGGAGGACAGGCCTCCGCCCTGCG  
GCGCCGGCACGGCAGAGGACATTGTGGCACTGGCACGGCAGAGAACTGTGGCACCGGC  
GGGCGCGGACGTTCCAGGTGGGCACTCCAGCCACCTGGGGAGTGGGCGAGGGTCCGAG  
GGCACTCTCCCTCAGCTCTCATCCCCGTGCCCCAGGTGGGAGGACGGCGCGCGCGC  
**A**ACTTTGAATCATCCGTGTGAAGAGGGCTGCGGCTTCCCCGGTCCCAAATGAAGGGGCGG  
**TTCCCGGACCCCTGTTTGCTTTAGAGCCGAGCGCGCCGATGCCCTCACACTCTGCGCCT**  
**CCTCTCCCGGACTTTAACACCCCGCTCTCTCCACCGACAGGTGACCTTTTGCTCTGA**  
**GCCAGATCAGAGAAGAAATTCGGTGCCGTGCGGGACG**ATGCCTGAGCAGGAGAGACAGAT  
CACAGCCAGAGAAGGGGCCAGTCCGAAA...

MREa

MREe

chr13:g.52,586,149T>C

MREc

MREd

5'UTR

coding region

Primer binding sites: **ATP7B K3.2 Fo** and **ATP7B K3.2 Re**

**Fig.S1: Promoter fragment cloned by us for luciferase reporter assay and localization of MREs (as described by Oh et al.<sup>1</sup>) as well as the variant g.52,586,149T>C (NC\_000013.10, hg19) detected and investigated by Chen et al.<sup>2</sup>**

[1] Oh, W. J., Kim, E. K., Park, K. D., Hahn, S. H. & Yoo, O. J., Cloning and characterization of the promoter region of the Wilson disease gene, *Biochem. Biophys. Res. Commun.* **259**, 206-211 (1999).

[2] Chen, H., Jagadeesh, K., Birgmeier, J., Wenger, A., Guturu, H., Schelley, S. et al., An MTF1 binding site disrupted by a homozygous variant in the promoter of *ATP7B* likely causes Wilson Disease., *Eur. J. Hum. Genet.* (2018).