# Higher HOPX expression is associated with distinct clinical and biological features and predicts poor prognosis in *de novo* acute myeloid leukemia

**Chien-Chin Lin,[1,2,3] Yueh-Chwen Hsu,[3] Yi-Hung Li,[2] Yuan-Yeh Kuo,[4] Hsin-An Hou,[2] Keng-Hsueh Lan,[5] Tsung-Chih Chen,[2] Yi-Shiuan Tzeng,[4] Yi-Yi Kuo,[2] Chein-Jun Kao,[2] Po-Han Chuang,[2] Mei-Hsuan Tseng,[2] Yu-Chiao Chiu,[6] Wen-Chien Chou[1,2] and Hwei-Fang Tien[2]**

[1]Department of Laboratory Medicine; [2]Division of Hematology and Department of Internal Medicine; [3]Graduate Institute of Clinical Medicine; [4]Graduate Institute of Oncology, College of Medicine; [5]Division of Radiation Oncology and Department of Oncology and [6]Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei, Taiwan

 doi:10.3324/haematol.2016.161257

Received: November 30, 2016.

Accepted: March 17, 2017.

Pre-published: March 24, 2017.

Correspondence: wchou@ntu.edu.tw or f99945006@ntu.edu.tw or hftien@ntu.edu.tw

**Supplemental methods**

**Gene signature analysis**

GSEA tests whether content genes of a stem cell signature are overrepresented at either side of the global gene list ranked by *HOPX*-associated differential expression. Here we ranked genes based on their *t*-test *P* values between samples with high (top quartile) and low (bottom quartile) *HOPX* expression. The statistical significance of a GSEA enrichment score was tested against 2000-time permutations.

**Expression of *HOPX* isoforms**

Total RNA from the 56 prospectively recruited patients was isolated using TRIzol reagent and reverse-transcribed with the Goscript$^{TM}$ kit (Promega). The primer sequences and their locations for the isoforms of human *HOPX* are shown in Supplementary Table 1 and Supplementary Figure 1.

**Bisulfite treatment and methylation analysis of *HOPX***

The primers for bisulfite PCR cover dense CpG islands. Fifty ng of bisulfite treated DNA, 200 µM dNTP mix, 0.2 U of Hot Start Taq DNA polymerase, 200 nM forward (aggaagagagGTGTGTAGTTTTGTTTGGAGAGGG) and reverse (cagtaatacgactcactatagggagaaggctATAATTTTAAAAAATCCCCTTAAAAACTTC) primers were mixed together with double-distilled water to a total volume of 5µl. The *in vitro* transcription and base specific cleavage reaction were performed with MassCLEAVE kit (Sequenom, San Diego, CA). The methylation level was analyzed by

EpiTYPER software.

**Statistical analysis**

All statistical analyses were performed using XLSTAT statistical analysis software edition 2015.1 (Addinsoft, Deutschland, Germany). Whole patient population (n = 347) were included for analyses of the correlation between mRNA expression and clinical characteristics, but only those 227 patients who received standard chemotherapy were included in analyses of survivals.

**Supplementary Table 1.**

**Primer sequences of the five human *HOPX* isoforms**

| | Forward 5’ to 3’ | Reverse 5’ to 3’ |
|---|---|---|
| NM_032495 (a), NM_139212 (b1), and NM_001145460 (c) | TCGAGTGTGTGCTCATAGGC | TTTGGAAGCTGTGTTTGCTG |
| NM_139211 (b2) | TTAGAGCCGGAGCGCGCA | GTGCTTGTCGACCTTGTTGA |
| NM_001145459 (b3) | CTTCCTTAGAGCCGGAGGTC | GTGCTTGTCGACCTTGTTGA |

**Supplementary Table 2.**

**Comparison of clinical manifestations between AML patients receiving standard chemotherapy with higher and lower *HOPX* expression**

| Variables | Total (n=227) | Higher *HOPX* Expression (n=103) | Lower *HOPX* Expression (n=124) | *P* value |
|---|---|---|---|---|
| **Sex**[†] | | | | 0.681 |
| Male | 118 | 52 | 66 | |
| Female | 109 | 51 | 58 | |
| **Age (year)**[‡] | | 47 (15-76) | 46 (18-84) | 0.716 |
| **Lab data**[‡] | | | | |
| WBC (/μL) | | 24110 (580-341420) | 23425 (380-423000) | 0.920 |
| Hb (g/dL) | | 8.3 (3.3-13.0) | 7.9 (3.7-16.2) | 0.826 |
| Platelet (×1,000 /μL) | | 65.0 (6-655) | 38.5 (2-412) | 0.002 |
| Blast (/μL) | | 12380 (0-260615) | 9802 (0-348777) | 0.262 |
| LDH (U/L) | | 849 (202-7734) | 1035 (242-13130) | 0.014 |
| **FAB**[*] | | | | <0.001 |
| M0 | 2 | 2 (83.3) | 0 (16.7) | 0.119 |
| M1 | 55 | 33 (62.7) | 22 (37.3) | 0.012 |
| M2 | 73 | 28 (44.0) | 45 (56.0) | 0.144 |
| M3 | 26 | 3 (14.3) | 23 (85.7) | <0.001 |
| M4 | 55 | 31 (56.3) | 24 (43.7) | 0.060 |
| M5 | 12 | 2 (20.0) | 10 (80.0) | 0.040 |
| M6 | 4 | 4 (87.5) | 0 (12.5) | 0.027 |

[†]number of patients

[‡]median (range)

*number of patients (% with higher or lower *HOPX* expression in the AML subtype)

Abbreviation: LDH, lactate dehydrogenase; CR, complete remission; PR, partial remission

**Supplementary Table 3.**

**Association of *HOPX* expression levels with cytogenetic abnormalities**

| Variables | Total | Higher *HOPX* Expression | Lower *HOPX* Expression | *P* |
|---|---|---|---|---|
| **Karyotype†** | | | | |
| Favorable | 58 | 11 | 47 | <0.001 |
| t(8;21) | 24 | 0 | 24 | <0.001 |
| t(15;17) | 27 | 4 | 23 | <0.001 |
| Intermediate | 196 | 102 | 94 | 0.532 |
| Normal | 166 | 81 | 85 | 0.582 |
| Unfavorable | 71 | 48 | 23 | 0.001 |

†Favorable, t(15;17), t(8;21), inv (16); unfavorable, -7, del(7q), -5, del(5q), 3q abnormality, complex abnormalities; Intermediate, normal karyotype and other abnormalities.

**Supplementary Table 4.**

**Univariate analysis on overall survival**

| Variables | Overall Survival | |
|---|---|---|
| | Months[#] | P |
| ***NPM1*$^{+}$/ *FLT3*-ITD$^{-}$** | | 0.197 |
| Yes (n=29) | NR | |
| Others (n=198) | 48.8 (36.3-61.3) | |
| ***CEBPA*** | | 0.002 |
| Double mutation (n=25) | NR | |
| Others (n=202) | 39.2 (28.9-49.5) | |
| ***FLT3*-ITD** | | 0.001 |
| Mutated (n=63) | 18.0 (12.3-23.7) | |
| Wild (n=164) | 108.1 (82.5-133.7) | |
| ***FLT3*-TKD** | | 0.828 |
| Mutated (n=22) | 39.2 (12.3-66.1) | |
| Wild (n=205) | 50.0 (38.8-61.2) | |
| ***RUNX1*** | | 0.040 |
| Mutated (n=24) | 24.9 (17.0-32.7) | |
| Wild (n=203) | 59.3 (37.5-81.0) | |
| ***WT1*** | | 0.020 |
| Mutated (n=23) | 14.7 (12.4-17.0) | |
| Wild (n=204) | 59.3 (47.8-70.9) | |
| ***IDH2*** | | 0.430 |
| Mutated (n=27) | 66.0 (41.4-90.6) | |
| Wild (n=200) | 50.0 (37.9-62.1) | |
| ***ASXL1*** | | 0.815 |
| Mutated (n=19) | 22.0 | |
| Wild (n=208) | 54.4 (43.2-65.6) | |
| ***DNMT3A*** | | 0.313 |
| Mutated (n=40) | 39.2 (22.8-55.6) | |
| Wild (n=187) | 59.3 (46.2-72.4) | |
| ***IDH1*** | | 0.958 |
| Mutated (n=13) | 57.4 (24.4-90.4) | |
| Wild (n=214) | 50.0 (38.4-61.6) | |
| ***TET2*** | | 0.093 |
| Mutated (n=31) | 16.0 (9.4-22.6) | |
| Wild (n=196) | 57.4 (48.9-65.9) | |

| | | |
|---|---|---|
| ***PTPN11*** | | 0.898 |
| Mutated (n=12) | NR | |
| Wild (n=215) | 50.0 (38.3-61.7) | |
| ***NRAS*** | | 0.160 |
| Mutated (n=39) | 66.0 | |
| Wild (n=188) | 48.8 (37.7-59.9) | |
| ***KIT*** | | 0.350 |
| Mutated (n=13) | 17.4 (10.5-24.3) | |
| Wild (n=214) | 54.4 (42.2-66.4) | |
| ***KRAS*** | | 0.164 |
| Mutated (n=11) | 14.0 (0.1-27.9) | |
| Wild (n=216) | 54.4 (44.9-63.9) | |
| ***MLL*-PTD** | | 0.001 |
| Mutated (n=9) | 10.5 (6.3-14.7) | |
| Wild (n=217) | 57.4 (47.25-67.55) | |
| ***TP53*** | | <0.001 |
| Mutated (n=5) | 2.5 (1.3-3.7) | |
| Wild (n=222) | 57.4 (47.9-66.9) | |
| ***HOXA9*** | | <0.001 |
| Higher expression (n=114) | 24.9 (19.3-30.5) | |
| Lower expression (n=113) | 116.8 | |
| ***HOPX*** | | <0.001 |
| Higher expression (n=103) | 23.7 (18.2-29.2) | |
| Lower expression (n=124) | 116.8 | |

Abbreviation: NR, not reached

[#] Months: median±S.D.

**Supplementary Table 5.**

**Common leading-edge genes of the HSC and LSC signature among NTUH, TCGA, GSE12417 datasets**

| HSC signature | LSC signature |
|---|---|
| *ABCB1* | *ATP1B1* |
| *BAALC* | *C2CD2* |
| *BCL11A* | *FLJ13197* |
| *C5ORF23* | *IQGAP2* |
| *CRIM1* | *TGIF2* |
| *DAPK1* | |
| *GUCY1A3* | |
| *HTR1F* | |
| *INPP4B* | |
| *KIAA0125* | |
| *MLLT3* | |
| *MYO5C* | |
| *PLSCR4* | |
| *PROM1* | |
| *SOCS2* | |
| *SPINK2* | |
| *TFPI* | |

**Supplementary Table 6.**

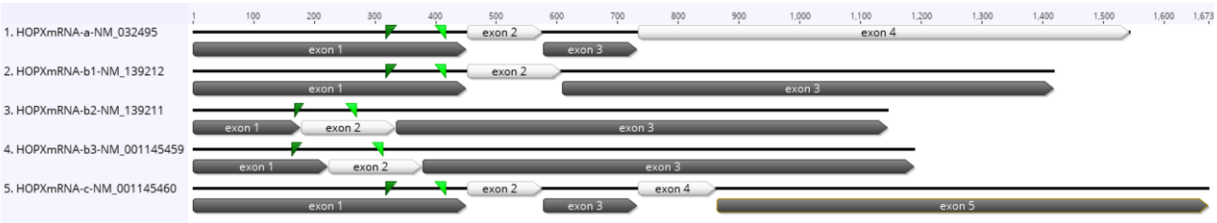
**Association between the expression levels of *HOPX* and the ABC genes related to chemoresistance**

| | Probe | Higher *HOPX* | Lower *HOPX* | *P* value |
|---|---|---|---|---|
| ***ABCB1*** | **4210039** | 5.879 | 5.455 | < 0.001 |
| ***ABCB1*** | **1230048** | 6.557 | 5.946 | < 0.001 |
| ***ABCG1*** | **6060377** | 7.611 | 6.880 | < 0.001 |
| ***ABCG1*** | **5860377** | 6.626 | 6.017 | < 0.001 |
| ***ABCG1*** | **6450059** | 5.931 | 5.864 | < 0.001 |
| ***ABCG2*** | **7160220** | 5.330 | 5.274 | 0.001 |

**Supplementary figures**

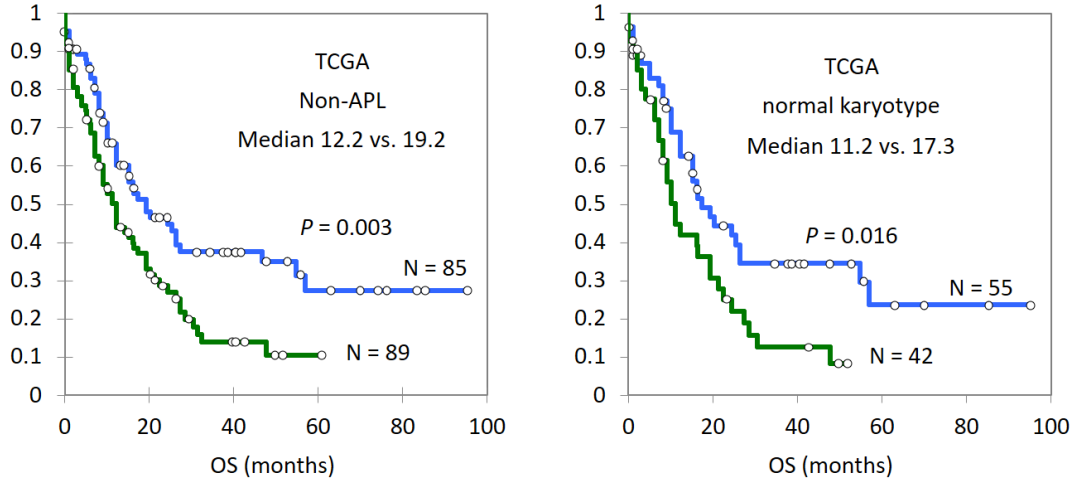
**Supplementary Figure 1**

Five isoforms of *HOPX*: *HOPXa* (NM_032495), *HOPXb1* (NM_139212), *HOPXb2* (NM_139211), *HOPXb3* (NM_001145459), *HOPXC* (NM_001145460) and the primer locations. The isoforms of a, b1, and c have low levels of expression and are summed up together by one pair of primers.



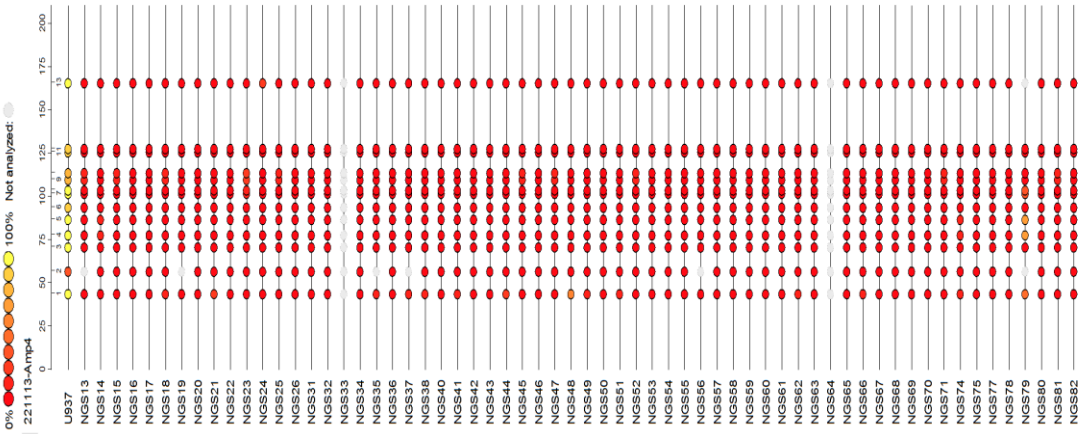

**Supplementary Figure 2**

Comparison of overall survival between patients with higher and lower *HOPX* levels in non-APL patients (A) and patients with normal karyotype (B) in the TCGA cohort. Green line: higher *HOPX* expression group; Blue line: lower *HOPX* expression group.



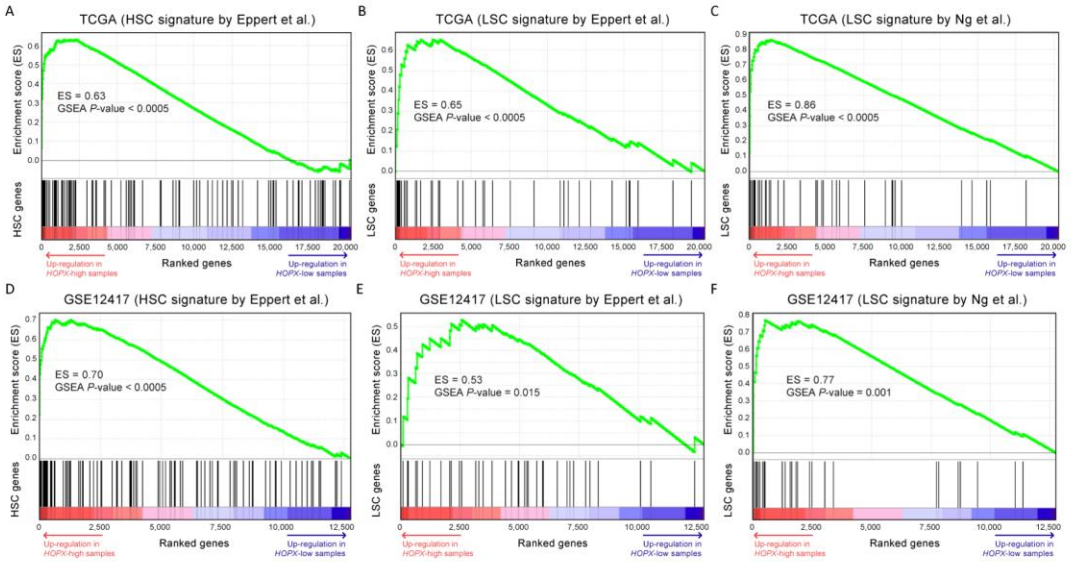
**Supplementary Figure 3**

*HOPX* b2 promoter CpG island methylation status in the U937 leukemia cell line and 62 AML patients. The red circles are not methylated while the yellow circles in U937 cells mean heavy methylation. The coordinate for the transcription start site is 0 (NM_139211).



**Supplementary Figure 4**

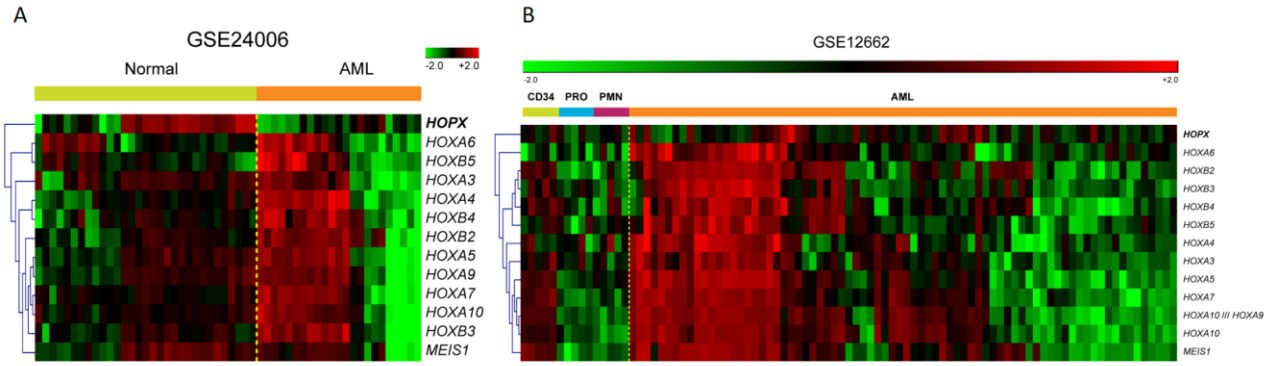
GSEA plots of HSC and LSC gene signatures in (A, B, C) TCGA and (D, E, F) GSE12417 datasets.




**Supplementary Figure 5**

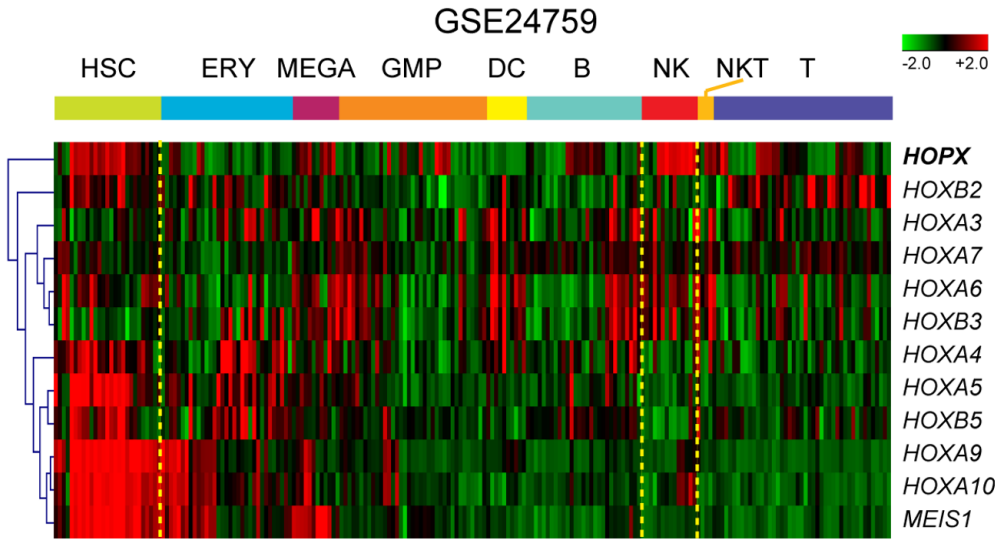
Comparison between expression patterns of *HOPX* and *HOX* family genes in normal/AML samples. (A) Heatmap of *HOPX*, *HOX* family genes, and ABC transporter genes in GSE24006. The dataset is composed of expression profiles of cells sorted from normal BM and umbilical cord blood (n=31) and from AML BM and peripheral blood (n=23), separated by a yellow line. The data obtained from probes representing the same gene are averaged; genes were presented with average-linkage hierarchical clustering. Samples are clustered in normal and AML groups, respectively. In normal samples, *HOPX* expression is generally concordant with *HOX* family genes (mean correlation coefficient, 0.37). In contrast, in the AML patients,

*HOPX* expression is generally discordant with *HOX* family genes (mean correlation coefficient, -0.31). (B) Heatmap of *HOPX*, other HOX genes, and ABC genes in GSE12662. The dataset is composed of CD34+ cells (CD34; n=5), promyelocytes (PRO; n=5), and neutrophils/polymorphonuclear leukocytes (PMN; n=5) fractionate from normal bone marrow, and 76 bone marrow samples of AML. The data from probes representing the same gene are averaged; genes are presented with average-linkage hierarchical clustering.



**Supplementary Figure 6**

Heatmap of GSE24759, which contains 9 distinct hematopoietic cell populations. The dataset profiles cells purified from 211 normal umbilical cord blood and peripheral blood samples: Hematopoietic stem cells (HSC; n=27), erythroid cells (ERY; n=33), megakaryocytes (MEGA; n=12), granulocyte/monocyte progenitors (GMP; n=37), dendritic cells (DC; n=10), B cells (B; n=29), natural killer cells (NK; n=14), natural killer T cells (NKT; n=4), and T cells (T; n=45). All *HOPX* and *HOX* family genes are highly expressed in HSC (average *z*-values = 0.82 and 0.75; both *P*-values < 0.001, one-sample *t*-test against zero).



**Supplementary Figure 7**

Comparison of global methylation patterns between *HOPX* and *HOX* family genes. (A) Heatmap of the 194-sample TCGA methylation dataset. We represent methylation levels by *M*-values, with positive and negative values indicating high and low levels of methylation, respectively. Hierarchical clustering reveals three clusters of genes, among which *HOXA3*, *HOXA4*, *HOXA5*, and *HOXB3* are highly methylated, while methylation levels of *HOXA7*, *HOXA9*, and *HOXB4* are low. (B) Distribution curves of methylation *M*-values of genes representing the three clusters. There are distinct patterns of methylation among the homeobox genes.