Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Ji J, Shi J, Budhu A, et al. MicroRNA expression, survival, and response to interferon in liver cancer. N Engl J Med 2009;361:xxxx-xx.

Supplementary text

MicroRNA expression, survival, and response to interferon in liver cancer

Junfang Ji[1], Jiong Shi[1,4], Anuradha Budhu[1], Zhipeng Yu[1], Marshonna Forgues[1], Stephanie Roessler[1], Stefan Ambs[2], Yidong Chen[3], Paul S. Meltzer[3], Carlo M. Croce[5], Lun-Xiu Qin[4], Kwan Man[6], Chung-Mau Lo[6], Joyce Lee[7], Irene O.L. Ng[7], Jia Fan[4], Zhao-You Tang[4], Hui-Chuan Sun[4], Xin Wei Wang[1]

[1]Liver Carcinogenesis Section, [2]Breast and Prostate Unit, Laboratory of Human Carcinogenesis and [3]Genetics Branch, CCR, NCI, NIH, Bethesda, MD, USA; [4]Liver Cancer Institute and Zhongshan Hospital, Fudan University, Shanghai, China; [5]Comprehensive Cancer Center, Ohio State University, Columbus, OH, USA; [6]Department of Surgery and [7]Department of Pathology, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong, China.

SUPPLEMENTARY METHODS

RNA Isolation and Real-Time qRT-PCR Analysis

Total RNAs were extracted from frozen tissues of cohort 1 using standard TRIZOL (Invitrogen, Carlsbad, CA) methods, and from paraffin-embedded tissues of cohort 2 and cohort 3 using a MasterPure RNA Purification Kit (Epicenter, Madison, WI). The expression of mature microRNAs was measured using Taqman MicroRNA Assays specific for miR-26a and miR-26b after reverse transcription (Applied biosystems, Foster City, CA). All comparisons between strata (sex. miR etc) were within each cohort. The Taqman MicroRNA Assay for U6 RNA was used to normalize the relative abundance of microRNAs. The expression of IL-6 was measured using the Taqman Gene Assay specific for this gene after reverse transcription by using the High Capacity cDNA Archive Kit (Applied Biosystems, Foster City, CA). The Taqman gene assay for 18s was used to normalize the relative abundance of mRNA. The experiments were performed in triplicate.

Microarray analyses and Statistics

For microRNA microarray profiling, tumors and paired non-tumor tissues were profiled separately using a single channel array platform previously described (1). The array quality control, data preprocessing and normalization were done essentially as previously described. The BRB-ArrayTools software 3.6.0 (http://linus.nci.nih.gov) was used for microarray analyses as previously described (2;3). MicroRNA probes with values missing from more than 50% of the arrays and those with less than 20% of expression data values having at least a 1.5-fold change in either direction from the probe's median value were excluded from the analysis, which left 624 probes. Class comparison analysis using t-statistics was used to identify microRNAs that were differentially expressed in tumors or surrounding non-cancerous tissues between males and females. For this analysis, the initial significance threshold of univariate tests was set at $p<0.05$ and the analyses were based on 1000 permutations for the multivariate test to generate permutation p-values for the global test to control false discovery rates. For mRNA

expression microarray profiling, we used our previously available oligoarray dataset based on a dual-channel platform (i.e., T/NT ratio) (4) that contained 224 cases matched to those with available microRNA microarray data described above. We used the median expression in tumors to dichotomize HCC cases, where low miR-26 expression was classified as the lower 50th percentile and high miR-26 expression was classified as the upper 50th percentile. Class comparison analysis based on dichotomized miR-26 expression levels was used to identify differentially expressed mRNAs between low miR-26 and high miR-26 HCCs. The same probe filtering criteria was followed as described above, leaving 11,580 expression probes for these comparisons.

Six class prediction algorithms, i.e., Support Vector Machines (SVM), Compound Covariate Predictor (CCP), Diagonal Linear Discriminant (DLD), 1-Nearest Neighbor (1NN), 3-Nearest Neighbor (3NN) or Nearest Centroid (NC), were also used to determine whether mRNA expression patterns could accurately discriminate low miR-26 HCCs from high miR-26 HCCs. In these analyses, 90% of the samples were randomly chosen to build a classifier which was then used to predict the remaining 10% of the cases. The accuracy of the prediction was calculated after 1000 repetitions of this random partitioning process to control the number and proportion of false discoveries. Hierarchical clustering analysis was performed using BRBarrayTools with median-centered correlation and complete linkage. Using an unsupervised approach, we also performed multidimensional scaling analysis using all cohort 1 samples based on the first three principal components of 11,580 genes that passed the filter. The expression levels of these genes were log-transformed, and Euclidean distance was used to determine their positions. Gene Network Analyses were used to identify signaling pathways that were enriched with genes differentially expressed in tumors between miR-26 low and miR-26 high HCCs using Ingenuity Pathways Analysis (Ingenuity®, www.ingenuity.com).

Kaplan-Meier survival analysis was used to compare patient survival based on dichotomized miR-26 expression, using GraphPad Prism software 5.0 (GraphPad Software, San Diego, CA) with statistical *P* values generated by the Cox-Mantel log-rank test. Cox proportional hazards regression analyses were used to analyze the effect of clinical variables on patient survival using STATA 9.2 (College Station, TX). A univariate test was used to examine the influence of each clinical variable on survival. A multivariate analysis was performed considering clinical variables from the univariate analysis that were significantly associated with survival with significance set at $p<0.05$. Multi-colinearity of the covariates was assessed and was not found to be present. In the final models, sex was included as a covariate due to its biological relevance in HCC outcome and its association with miR-26 expression. It was determined that the final models met the proportional hazards assumption. For RT-PCR data, the statistical *P* value, generated by the student *t*-test, and the Spearman correlation constant were calculated using GraphPad Prism Software 5.0. The statistical significance was defined as $p < 0.05$. All p-values in this paper are two-sided.

SUPPLEMENTARY FIGURE LEGENDS

Suppl Figure 1. The abundance of miR-26 expression in male and female hepatic tissues and tumors from the test cohort. (A) The expression levels of miR-26a-2 in female (n=30) and male non-tumor hepatic tissues (n=194). (B) Comparisons of relative levels of miR-26a-2 between paired T and NT when dichotomized by miR-26 status. A median expression level was used as a cutoff. Low miR-26 expression was classified as the lower 50th percentile (with a mean 2.69-fold reduction in T compared to NT). High miR-26 expression was classified as the upper 50th percentile (with a mean 0.98-fold reduction in T compared to NT). The data in A and B were determined by microarray analysis and expressed as log 2 relative expression normalized to a disease-free normal liver pool (n=8). (C) miR-26a-2 expression levels in tumors and survival outcomes. (D-F) Similar results as in A-C with miR-26b expression status.

Suppl Figure 2. Sample stratification based on miR-26 expression status and its associated cellular genes. (A) HCC samples were classified based on their average microRNA expression (i.e., miR-26a-1, 26a-2, and 26b). Median expression was used to separate cases into low miR-26 (blue) and high miR-26 (red). The position of each dot (case) is determined by three microRNA expression levels (miR-26a-1, miR-26a-2 and miR-26b). The stratification outcomes were used to generate the MDS plot. (B) A Venn diagram of mRNAs coexpressed in low miR-26 HCCs.

Suppl Figure 3. Expression of *S100P* and *SLC2A6* in HCC and their correlation between microarray and qRT-PCR. Expression levels of *S100P* (A) and *SLC2A6* (B) in 10 HCCs with low-miR-26 level and 10 HCC cases with high miR-26 level determined by qRT-PCR (left panel). A linear regression and correlation among data from qRT-PCR versus microarray is shown with r (spearman) and p-value indicated (right panel). Expression status is shown as the tumor (T)/ non-tumor (NT) ratio.

Suppl Figure 4. Expression of miR-26 and IL-6 in HCC. (A) Expression levels of IL-6 in 82 paired tumors (T) and non-tumor tissues (NT) determined by qRT-PCR. Student's t-test was performed to examine the IL-6 differential expression between T and NT. (B-C) Correlation of expression levels between IL-6 and miR-26a (B) or miR-26b (C) in 82 paired tumors and non-tumor tissues determined by qRT-PCR. The data are shown as the T/NT ratio on a log2 scale.

Suppl Figure 5. The abundance of miR-26 expression in male and female hepatic tissues and tumors from a validation cohort. Decreased expression of miR-26a and miR-26b in tumors (right panels) with a more abundant expression in female than male non-tumor tissues (left panels) is validated in an independent validation cohort from a retrospective randomized clinical trial. Expression levels of miR-26a and miR-26b were measured by qRT-PCR. P values are from un-paired t-tests.

Suppl Figure 6. The association of miR-26b expression in tumors with survival prognosis in two prospective randomized control trials for IFN adjuvant therapy. (A, B) The association of miR-26b expression with overall survival in control cases from cohort 2 (A)

or cohort 3 (D). Cohort 2: high miR-26b cases, n=23; low miR-26b cases, n=36. Cohort 3: high miR-26b cases, n=21; low miR-26b cases, n=19. (C, D) The association of IFN adjuvant therapy with overall survival in HCC patients with low miR-26b expression. Cohort 2: IFN cases, n=22; control cases, n=36. Cohort 3: IFN cases, n=20; control cases, n=19. (E, F) The association of IFN adjuvant therapy with overall survival in HCC patients with high miR-26a expression. Cohort 2: IFN cases, n=37; control cases, n=23. Cohort 3: IFN cases, n=19; control cases, n=21.

Reference List

(1) Budhu A, Jia HL, Forgues M, Liu CG, Goldstein D, Lam A et al. Identification of metastasis-related microRNAs in hepatocellular carcinoma. Hepatology 2008; 47(3):897-907.

(2) Budhu A, Forgues M, Ye QH, Jia LH, He P, Zanetti KA et al. Prediction of venous metastases, recurrence and prognosis in hepatocellular carcinoma based on a unique immune response signature of the liver microenvironment. Cancer Cell 2006; 10(2):99-111.

(3) Ye QH, Qin LX, Forgues M, He P, Kim JW, Peng AC et al. Predicting hepatitis B virus-positive metastatic hepatocellular carcinomas using gene expression profiling and supervised machine learning. Nat Med 2003; 9(4):416-423.

(4) Yamashita T, Forgues M, Wang W, Kim JW, Ye Q, Jia H et al. EpCAM and alpha-fetoprotein expression defines novel prognostic subtypes of hepatocellular carcinoma. Cancer Res 2008; 68(5):1451-1461.

Supplementary Table 1. Clinical Characteristics of Two Independent Clinical Trails

Clinical variable	IFN Test (Cohort 2) (n=135)	IFN Validation (Cohort 3) (n=79)	*p* value*
Sex			
Female	14	14	
Male	111	65	0.21[a]
No data	10	0	
Age-year			
Median (range)	50 (20-77)	52 (24-75)	0.23[b]
Alanine transaminase (ALT)			
Normal (≤50U/L)	107	44	
Abnormal (>50U/L)	16	35	<0.001[a]
No data	12	0	
HBV			
Negative	6	9	
Positive	118	70	0.10[a]
No data	11	0	
Tumor size-cm			
<=3	46	22	
>3	78	57	0.22[a]
No data	11	0	
Multinodular			
No	107	63	
Yes	17	16	0.25[a]
No data	11	0	
Cirrhosis			
No	16	0	
Yes	108	0	NA[c]
No data	11	79	
TNM Stage			
I	81	7	
II	29	34	
III-IV	14	38	<0.001[d]
No data	11	0	
Alpha fetoprotein (AFP)			
Negative (≤20ng/ml)	49	33	
Positive (>20ng/ml)	75	46	0.77[a]
No data	11	0	
IFNa Therapy			
Yes	72	39	
No	63	40	0.67[a]
Survival-month			
Median (range)	67 (2-82)	>60 (5-60)	0.06[e]

* [a]Fisher's exact test; [b]un-paired test; [c]Not available; [d]Chi-square test; [e]Log-rank test

Supplementary Table 2. Clinical Characteristic of Cases Used to Search for Sex-Related microRNAs

Clinical variable	Female (n=30) Value[a]	Male G1 (n=31) Value	Male G2 (n=31) Value	p value[b]		
				Female vs Male G1	Female vs Male G2	Male G1 vs Male G2
Age-year						
Median	52	52	52			
Range	25-72	26-71	27-71	0.93[c]	0.93[c]	1.00[c]
ALT[d]						
Normal	24	20	18			
Abnormal	6	11	13	0.26	0.10	0.80
Tumor size-cm						
<=3	16	9	10			
>3	14	22	21	0.07	0.12	1.00
Multinodular						
No	25	27	26			
Yes	5	4	5	0.73	1.00	1.00
TNM Stage						
I	15	14	13			
II	11	7	9			
III	4	10	9	0.18[e]	0.33[e]	0.84[e]
AFP[f]						
Negative	9	14	6			
Positive	20	17	25	0.30	0.38	0.06
No data	1	0	0			
Survival Months						
Range	3.0-67.1	4.5-67.1	2.3-65.4	0.31[g]	0.09[g]	0.55[g]

[a] Each value represents the number of patients.
[b] Fisher's exact test
[c] Un-paired t-test
[d] Normal: ≤50 (U/L); Abnormal >50 (U/L)
[e] Chi-square test
[f] Negative: ≤20 (ng/ml); Positive >20 (ng/ml)
[g] Log-rank test

Supplementary Table 3. Eight Sex-Related microRNAs

Sex-related miRNA	Genomic Location	Parametric p-value	Permutation p-value	Mean Intensities in Female	Mean Intensities in Male	Expression in Female
Non-HCC						
miR-321[a]	-	0.001	0.003	4585	2618	up
miR-26a-1	3p22.3	0.01	0.01	19879	14417	up
miR-10b	2q31.1	0.02	0.02	735	535	up
miR-125b-1	11q24.1	0.02	0.02	4211	2838	up
miR-99b	19q13.41	0.04	0.04	2628	2071	up
miR-325	Xq21.1	0.05	0.06	1885	1146	up
miR-342	14q32.2	0.04	0.03	306	373	down
HCC						
miR-129-2	11p11.2	0.007	0.004	893	1179	down

[a] miR-321 is reported as a fragment of Arg-tRNA.

Supplementary Table 4. Top 20 list of 27 gene networks from Ingenuity Pathway Analysis

ID	Genes in Network	Score*	# of genes	Top Functions
1	14-3-3**, ACLY, CPOX, CTDP1,** Cyclin B**, DAXX, DNM1, GAPDH, GML, HK3, HUWE1, IRF5,** Jnk dimer**, LRDD, MED22, MED28,** Ndpk**, NELF, NME1, NME2, PDCD2, PIN1, PKMYT1,** RNA polymerase II**, RPS6KA1,** Rsk**, SFN, SNCAIP, STRAP, TBXAS1, TP53, TP53BP1, TTC5, ZBTB17, ZNF74**	42	29	Cancer, Cell-To-Cell Signaling and Interaction, Cellular Function and Maintenance
2	**AP3D1, ARRB1, ATP5E, ATP5I, ATP6V1D,** Calmodulin**, CCT6A, CD3EAP, CDC34,** Ck2**, DIRAS3, ENO1,** F Actin**, G6PD,**H+-transporting two-sector ATPase**, IL1A, INA,** Insulin**, MAP6, MRPS10,** Peptidylprolyl isomerase**, PPIA, PTPRZ1,** Ras homolog**, RHOT2, RPS2, SFRS1, SFRS11, SHROOM3, SNRP70, STK11, SUGT1, TALDO1, UBA52, WNT2**	40	28	Cancer, Cell Cycle, Carbohydrate Metabolism
3	**BCL3, BCR, BRE, C5AR1, CCL8, CDH22, DIABLO, ERC1,** Glutathione peroxidase**, GMFB, GRB7,** Ikb, IKK, IL1**, IL17A, LDHA, LY96, MTPN, NFATC2,** NFkB**, NOL14, NUP62, PNPT1, RBCK1, RIPK3, S100P, SAP30, SLC2A6,** Sod, Tnf receptor**, TNFRSF14, TRAF1, TRAF2, TRIB3**	38	27	Cell Death, Gene Expression, Carbohydrate Metabolism
4	Actin, Adaptor protein 2**, ALG5, ATP13A2, ATPase, BCAR1, BHLHB3, BRUNOL4,** Caspase**, CENTD2,** Cyclin A**, DDX11, DDX39,** E2f**, EID1, GARNL1, HIST1H2AG,** Histone h3**, HNRPK, KEAP1, KHDRBS1, MCM2, MCM5, PFDN4, PFDN5,** PI3K, Rb**, RUVBL1, RUVBL2, SYNCRIP, SYT3, TCF3, TGM2, THOC1, THOC2**	36	26	Gene Expression, RNA Post-Transcriptional Modification, Molecular Transport
5	Akt**, ALDOA, ATXN2L,** Cbp/p300**, CHFR, DGKZ, EPO,** Esr1-estrogen**, ETV6, FAM14A, FAM89B,** GC-GCR dimer**, HMGB1, IL12RB2, INPP5E,** JAK**, MPL,** N-cor**, NCOR2, NOSIP, NR0B1,** Nuclear factor 1**, OGG1, PGR, PRMT2, PTPRF, PTPRS (includes EG:5802),** STAT**, STAT2,** STAT5a/b, Thyroid hormone receptor**, TYK2, TYRO3, UCN, WDR1**	34	25	Cellular Growth and Proliferation, Immune and Lymphatic System Development and Function, Tissue Morphology
6	**ADM,** Alkaline Phosphatase, Ap1**, ARHGEF6, BIN1, BMP2K, CAD, CCL4, CD37, CD209, CSNK1E,** Dynamin**, ECE1, ELK1, ETV5,** Fgf**, FGF4, FGF12, J**UN/JUNB/JUND, LDL, Mapk**, MAX,** MHC Class II**, MSC, MXD3, NPC2,** PLC gamma**, PSMC3IP, SCARB2, SLA,SPRY1, SYK,** Tgf beta, VAV**, WIPF1**	32	24	Cellular Function and Maintenance,Cellular Compromise,Immune and Lymphatic System Development and Function

7	**B2M, BCL7B,** CD3, DUB**, EEF2, GPR109B,** Gsk3, Hexokinase**, HK1,** Ige**, IGHG1, IL10, JUN, KLRC3, MAFK,** Mek, MHC Class I**, MYB, Nfat, NMB,** Rap1, Ras**, RASSF5, RELB,RIT1, SAMD4A, SCRIB, SIT1,** Sos**, SPI1,** TCR**, TYROBP, USP11, USP22, USP33**	30	23	Immune and Lymphatic System Development and Function, Tissue Morphology, Developmental Disorder
8	**ADA,** Adenylate Cyclase**, ARF5, CACNG2, CALCB,** Calpain, CaMKII**, CAPN3, CAPN10, CORO1B,** Creb, ERK1/2, G-protein beta**, GLMN, GNAO1, GNB5, GRM3, HSF2,** Integrin**, ITGAM, KPNB1, MAG, MMP1,** PDGF BB, Pka, Pkc(s), PLC, Pld, PP2A**, PPP2R5B, RPS6KB1, SLC32A1, SNAP23, SNRPA, SYNE1**	28	22	Amino Acid Metabolism,Cell Cycle,Cell-To-Cell Signaling and Interaction
9	**ADRB1,** Alcohol group acceptor phosphotransferase**, AMPK,** Calcineurin protein(s**), CUL5,** DNA-directed RNA polymerase**, EWSR1, FZR1, GRK4, GTF2A2, HCFC1,** Hsp70, Hsp90, Jnk**, LDB3, LIMK1,** Mek1/2**,MYOZ3,** Nos**, NOS1,** P38 MAPK, p70 S6k, Pak**, PAK2,** Pdgf**, POLR1D, POLR2A, POLR2E,** Rac**, RRM2, SETD1A,** TFIIA**, TK1, TOM1,** Ubiquitin	22	19	Behavior,Amino Acid Metabolism, Cancer
10	**ADCYAP1, ANKRD11, BET1L,** beta-estradiol, CPLX2, CREB1**, ESRRA, FAM105A, FHL5, GALNT7, GOSR1, HEXIM1, HEXIM2,** NADH2 dehydrogenase (ubiquinone)**, NAPB,** NAPG, NCOA1, NCOA2, NDUFA3**, NDUFA7, NDUFA10,** NDUFC1, NDUFC2, NDUFS6, NDUFV3, phosphate**, PRPF31,** REST, SNAP25, Snare**, TRIM9, TSPAN14,** UHRF1BP1**, UQCR**	17	16	Gene Expression, Cell Death,Connective Tissue Development and Function
11	**ACPP,** beta-estradiol**, C11ORF10, CFD,** CHST3, CHST12, CHST13**, GSTM3,** HES1, HS3ST2, HS3ST5, HS3ST6**, HS3ST3A1,** HS3ST3B1, HS3ST4**, LAGE3, LHFPL2, MMD, MX2, NUDT1, PFKL, PPRC1, RBM15, SAPS2,** SMP2A, sulfotransferase, SULT1A2, SULT1A4, SULT1B1, SULT1C2, SULT1C3, SULT1C4, SULT4A1**, TMEM37,** UST	16	15	Carbohydrate Metabolism, Small Molecule Biochemistry, Amino Acid Metabolism
12	**ANAPC11, AOAH,** BUB1, DNAJA2, DNAJB1**, DNAJB5,** ENC1**, FKBP15, GBP5,** GCLM, HSPA9, hydrogen peroxide, IFI6, IFNB1**, LILRA2, LILRB3, LRRC47, MRPL20,** NFE2L2, OSGIN1**, PFDN5,** PRDX6, PSMD**, PSMD1, PSMD2,** PSMD5, PSMD7, PSMD9, PSMD12**, RAE1, RPL5,** SNCA, UQCRFS1**, UQCRH,** XAF1	16	15	Cancer,Cell Death,Cellular Compromise
13	ARHGEF2**, C20ORF117,** CDC45L**, CDCA7L,** CTSL2, CYFIP2**, EXOSC1,** EXOSC2, EXOSC3, EXOSC4, EXOSC5, EXOSC7, EXOSC8, EXOSC9**, FXC1,** GAS1, IFI202B, INSR, MAP4, MARK4, MNT, MXI1**, MYBBP1A,** MYC**, NOL5A, NUDC, PLEKHF1,** PRL2C2**, RPL32, RPS13, RPS20, RPS15A, STRA13, SURF6,** XRN1	14	14	RNA Post-Transcriptional Modification, Cell Cycle, Connective Tissue Development and Function

14	5-hydroxytryptamine, 5-hydroxytryptophan**, APBB2, CAMK1, CIB2,** DNAJC3, FKBP2, FKBP7, FKBP10, FKBP11**, FKBP14,** FKBP1B, GORASP2, GPNMB, HNF1A**, HNMT,** IL6, MIA2, PBSN, PDIA2, PIM1, PPIB, PRL**, RAB27B, RAB33B,** RABAC1**, REG1A, RP9, SBNO2, SLC4A2, SPCS3, TMED10, TUBA1B,** UNC13D, XBP1	14	14	Molecular Transport,Small Molecule Biochemistry,Cancer
15	**ACOT7,** APEX1, ARL4A, ARNT2**, CALB2,** CCNG2, CCS**, CHD2, CNDP2, FAM50A,** FKBP3, G6PD2**, GAK, GREM2,** GSTA3, HGF, HGFAC, HIF1A, IGF2, IL13RA2, IRS2, KLK1B9, KLK1B22**, KRT5, PCSK4,** PNRC1, progesterone**, RAB20, REG1B,** RNF103**, SMUG1,** SPSB1, STAT3**, TSSK2, ZNF592**	14	14	Cancer, Cellular Movement, Gastrointestinal Disease
16	**ADAMTS7,** ARSB, ARSC2, ARSD, ARSE, ARSF, ARSG, ARSH, ARSI, ARSJ, ARSK, Aryl Sulfatase**, B4GALT3,** CDKN2A**, GCM1,** GTSE1, HRAS**, ING1,** MIF**, NANOG, NSUN5C,** OTP, RECQL4**, RPL39,** S100B, SELPLG**, SULF2, TIMM13,** TOPORS**, TP53, TPP1, TTC1,** ZMAT3, Zn2+**, ZNF408**	14	14	Cellular Compromise, Cancer, Tumor Morphology
17	AP2A1, ARL4C, ATP, BATF3**, BCKDK, C14ORF153, C9ORF86,** CASP3, CD40LG, CEBPA**, CHKB,** DDX21, Ethanolamine kinase**, ETNK1, GRSF1,** IL2**, INCENP, KIF2A, KLRB1, LST1,** MAGEA3**, MRPS12,** MT1H, MT2A, PDCD1LG2, RNASE3**, SERPINB1,** SPINK7, SWAP70**, TBCD,** TNFRSF13C, TUBA3C, TUBB2A**, TUBG1**	14	14	Cancer, Infectious Disease, Cellular Growth and Proliferation
18	amino acids, APP, BTK**, CDC2L2,** CHAT, CLSTN1**, CPM, DAPK3, DGUOK,** DPYSL2**, DUSP7,** DYNC1I1**, FANCC,** FER, FPRL1, FYB, HOMER1**, IBTK, ICMT,** KIF5A, KIF5B, KIF5C**, KLC1, KLC2,** LILRA6, MAPK8IP2, MAPK8IP3**, MPZL1, PHKG2, PHLDA2,** PPME1, PTPN11, SPTAN1, ST8SIA1, TXK	13	13	Amino Acid Metabolism, Post-Translational Modification, Small Molecule Biochemistry
19	ARHGAP10, ARHGAP26, BCLAF1, Ck2,**DEDD**, DEDD2, **DSG2**, EGF, EIF5, **EIF4A2**, EIF4E, EIF4G3, **GLE1**, GTP, **MYCBPAP**, NOL3, NUP155, NUPL2, **PLEKHG5**, POP1, POP4, **POP7**, PXN, **RAB11B, REPS1**, RHOA, **RHPN2**, RND2, **RPP21**, RPP30, RPP38, RPP40, **SAFB2**, TCOF1, **UBOX5**	13	13	Protein Synthesis, Cell Signaling, DNA Replication, Recombination, and Repair
20	AATF, **BUB3**, BUB1, **C17ORF49**, CDKN1A, CHMP4A, CHMP4B, **CHMP4C**, CPSF3, CSTF2, **EAF1**, ELL, ELL2, F2, HSPA9, IL32, **LRSAM1**, MET, **MGRN1**, MLL, NCBP2, **NCBP1**, PDCD6IP, **SPSB2, SRM, STAMBP**, SYMPK, TSG101, **TUBG2**, UBE2S, VPS24, VPS28, **VPS37C**, VPS4A, **ZNF205**	13	13	RNA Post-Transcriptional Modification, Cellular Development, Hematological Disease

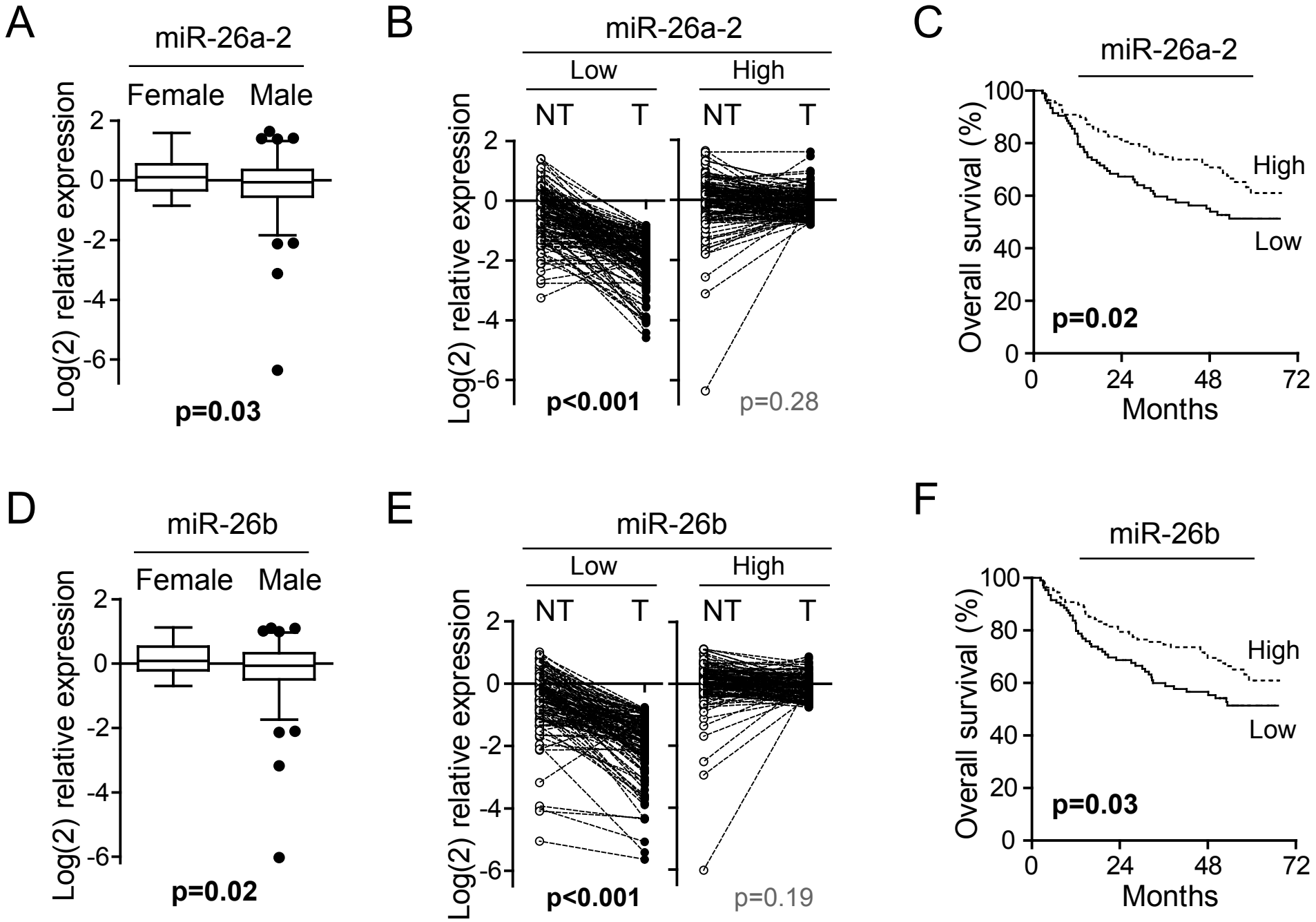
* The score is a numerical value used to rank networks according to how relevant they are to the genes in the input dataset (770 genes). The score takes into account the number of genes in the networks and the size of the network to approximate how relevant this network is to the input gene list. Gene symbols highlighted in bold are those included in the input gene list.

Supplementary Table 5. Univariate and Multivariate Cox Regression Analysis of miR-26 Expression Levels and Overall Survival in Subjects with HCC[a]

Clinical Variable	Hazard Ratio (95% CI[b])	p-value
UNIVARIATE ANALYSIS[c]		
miRNA-26a (Low vs. High)	2.3 (1.1-5.0)	**0.03**
miRNA-26b (Low vs. High)	2.3 (1.1-4.9)	**0.04**
Age (≤50 years vs. >50 years)	0.9 (0.5-1.5)	0.57
Sex (Male vs. Female)	1.4 (0.6-3.7)	0.47
AFP (>20ng/ml vs. ≤20ng/ml)	1.2 (0.6-2.5)	0.58
Cirrhosis (Yes vs. No)	0.7 (0.3-1.8)	0.46
ALT (>50U/L vs. ≤50U/L)	1.2 (0.5-2.8)	0.73
Tumor size (>3cm vs. ≤3cm)	1.2 (0.6-2.4)	0.56
Tumor encapsulation (No vs. Yes)	1.3 (0.7-2.7)	0.39
Multinodular (Yes vs. No)	1.1 (0.5-2.7)	0.81
TNM staging (II-III vs. I)	2.2 (1.2-4.3)	**0.02**
MULTIVARIATE ANALYSIS[d] for miR-26a		
miRNA-26 (Low vs. High)	2.2 (1.0-4.7)	**0.05**
TNM staging (II-III vs. I)	1.9 (1.0-3.9)	**0.05**
Sex (Male vs. Female)	1.1 (0.4-2.8)	0.88
MULTIVARIATE ANALYSIS[d] for miR-26b		
miRNA-26 (Low vs High)	2.2 (1.1-4.9)	**0.04**
TNM staging (II-III vs. I)	2.1 (1.1-4.2)	**0.03**
Sex (Male vs. Female)	1.0 (0.4-2.6)	0.99

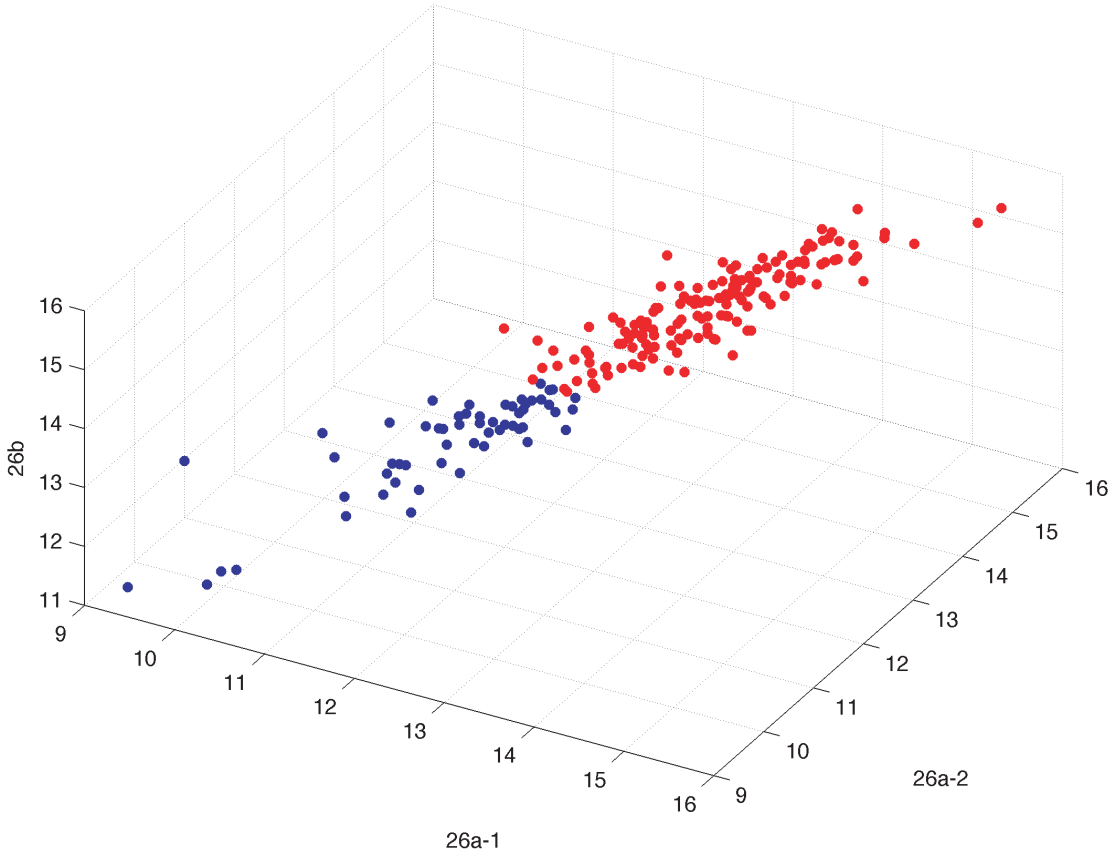
[a]The analysis was performed on control cases (n=60) (cohort 2) dichotomized by miR-26a/b low level group and miR-26a/b high level group; [b]95% CI, 95% confidence interval; [c]Univariate analysis, Cox proportional hazards regression; [d]Multivariate analysis, Cox proportional hazards regression; Significant p values (<0.05) are highlighted in bold.

Supplementary Fig 1

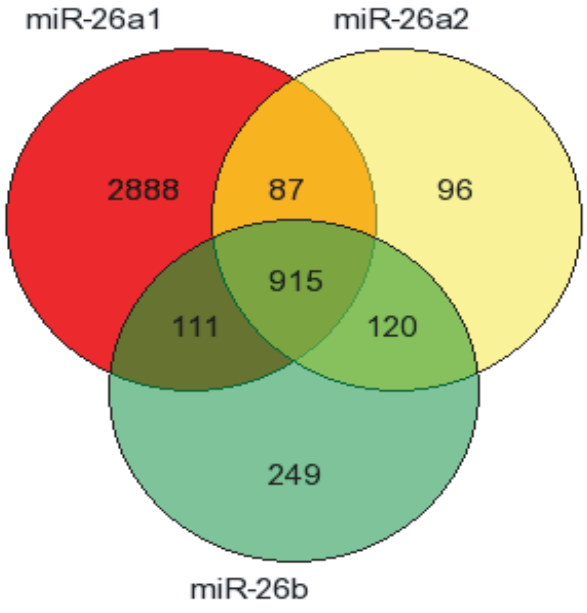


Supplementary Fig 2

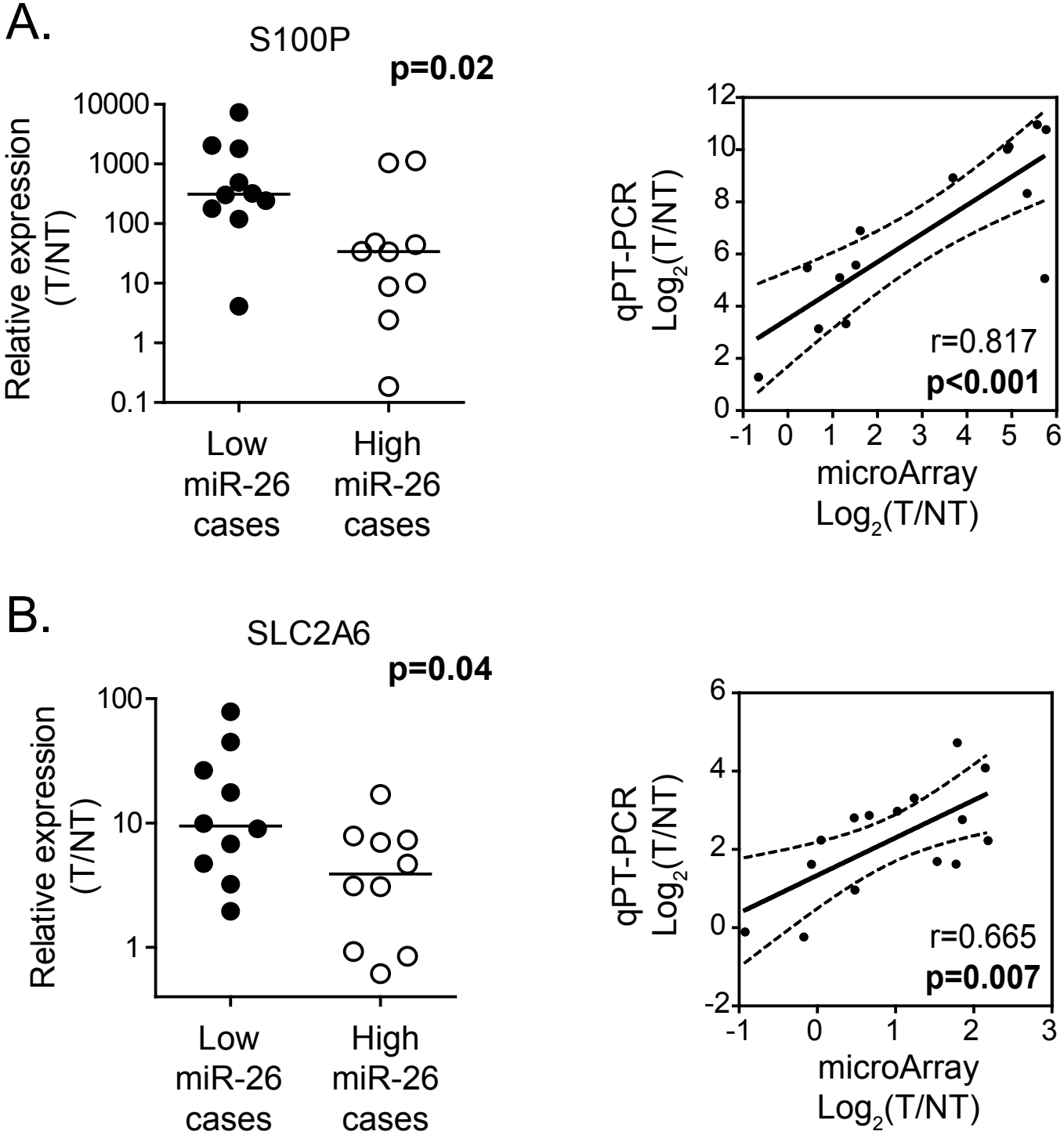
A



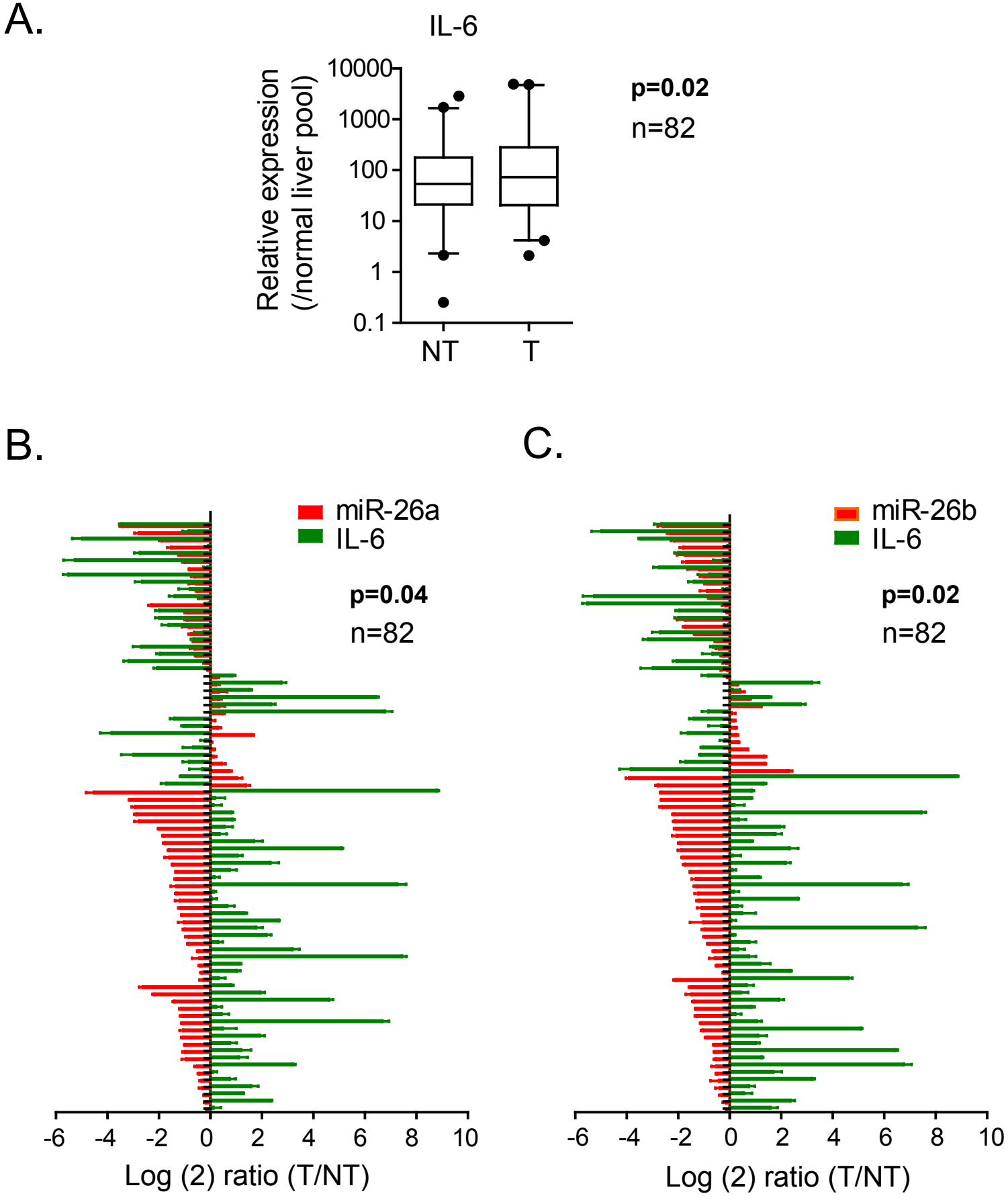
B



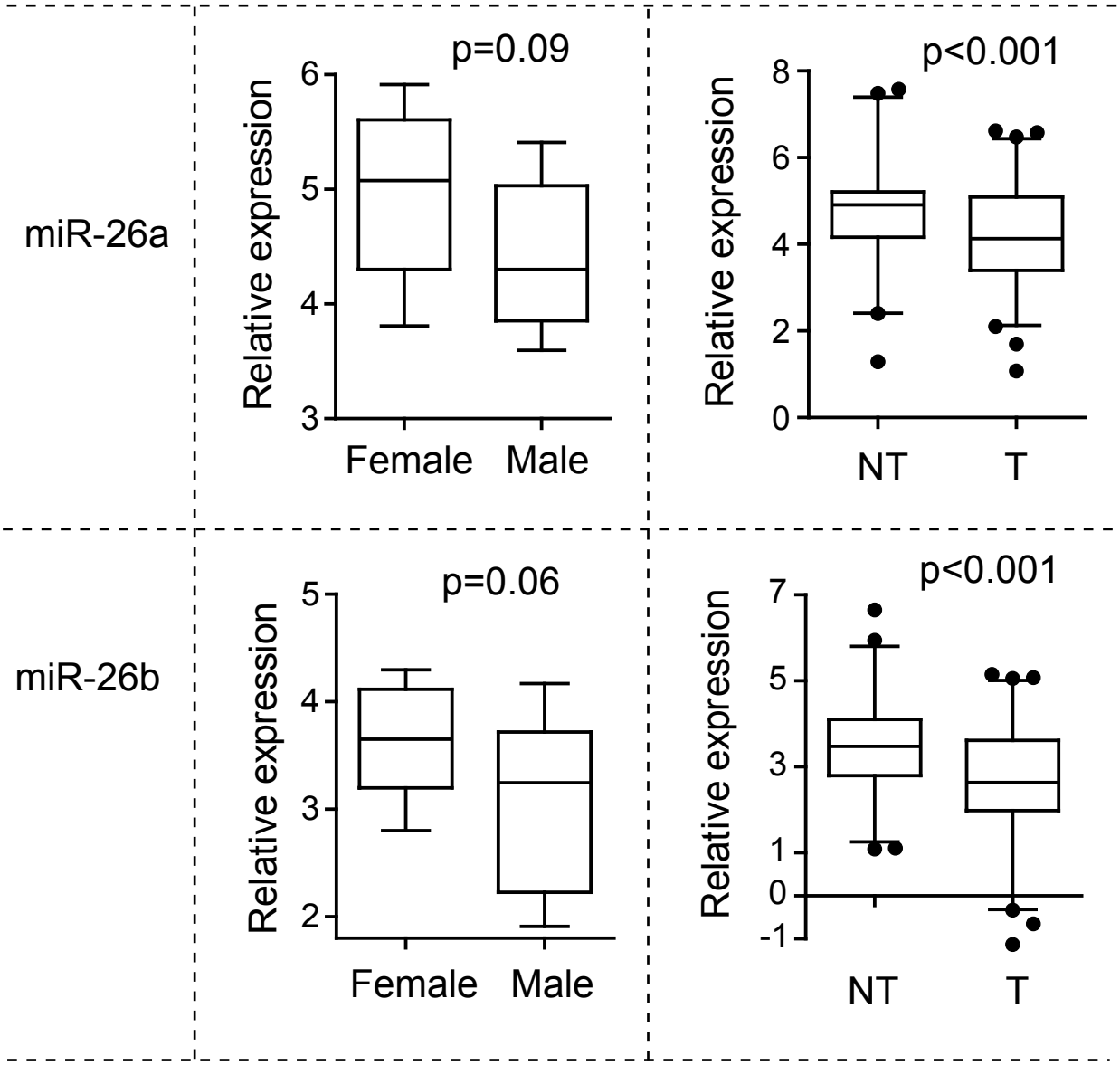
Supplementary Fig 3



Supplementary Fig 4



Supplementary Fig 5



Supplementary Fig 6