Supplementary materials

Genome-wide lncRNA microarray profiling identified novel circulating lncRNAs for detection of gastric cancer

Kecheng Zhang[1*], Hongzhi Shi[2*], Hongqing Xi[1*], Xiaosong Wu[1], Jianxin Cui[1], Yunhe Gao[1], Wenquan Liang[1], Chong Hu[1], Yi Liu[1], Jiyang Li[1], Ning Wang[1], Bo Wei[1], Lin Chen[1]

[1] Department of General Surgery & Institute of General Surgery, Chinese People's Liberation Army General Hospital, Fuxing Road 28, 100853 Beijing , P.R. China

[2] Department of General Surgery, General Hospital of Armed Police Force, Yongding Road 69, 100039 Beijing, P.R. China

[*] Contributed equally to this work

Figure S1. Association of blood cells (including white blood cells (WBCs), red blood cells (RBCs), platelets (PLTs) and hemoglobin (Hb)) with lncRNAs *TINCR* (A), *CCAT2* (B), AOC4P (C), *BANCR* (D) and *LINC00857* (E), respectively.

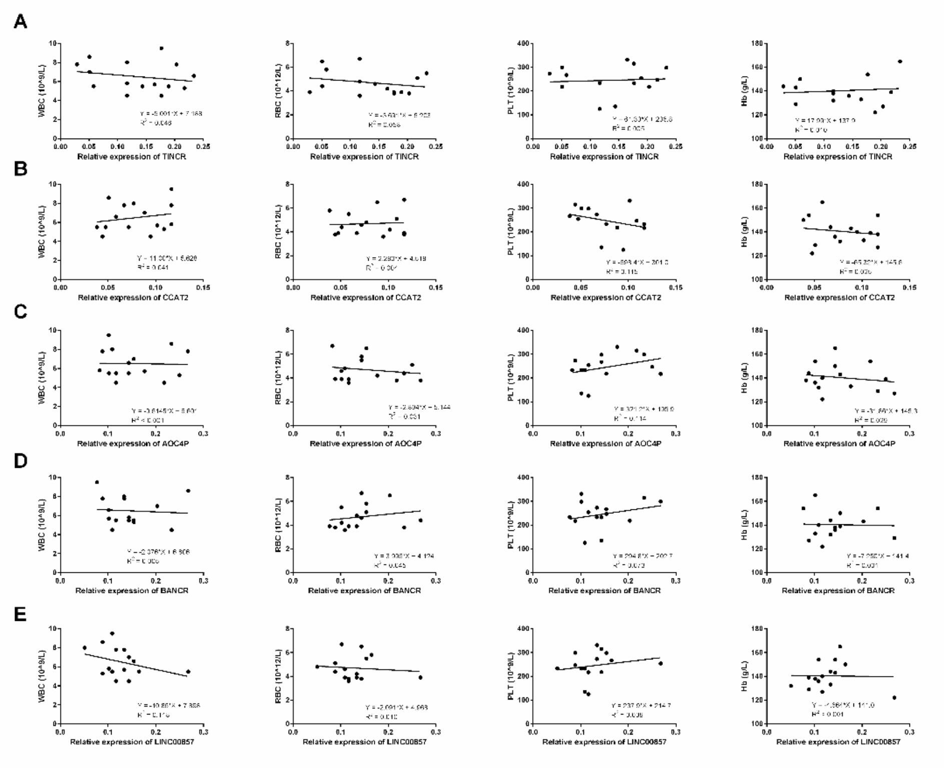

Figure S2. Box and whisker and ROC plots for individual lncRNAs *TINCR* (A), *CCAT2* (B), AOC4P (C), *BANCR* (D) and *LINC00857* (E) in the training phase. The horizontal line in the middle of each box indicates the median, whereas the top and bottom borders of the box mark the 95th and 5th percentiles, respectively. The whiskers above and below the box extend to the most extreme point no longer than 1.5 times the interquartile range from the box. The Mann-Whitney U test was used.

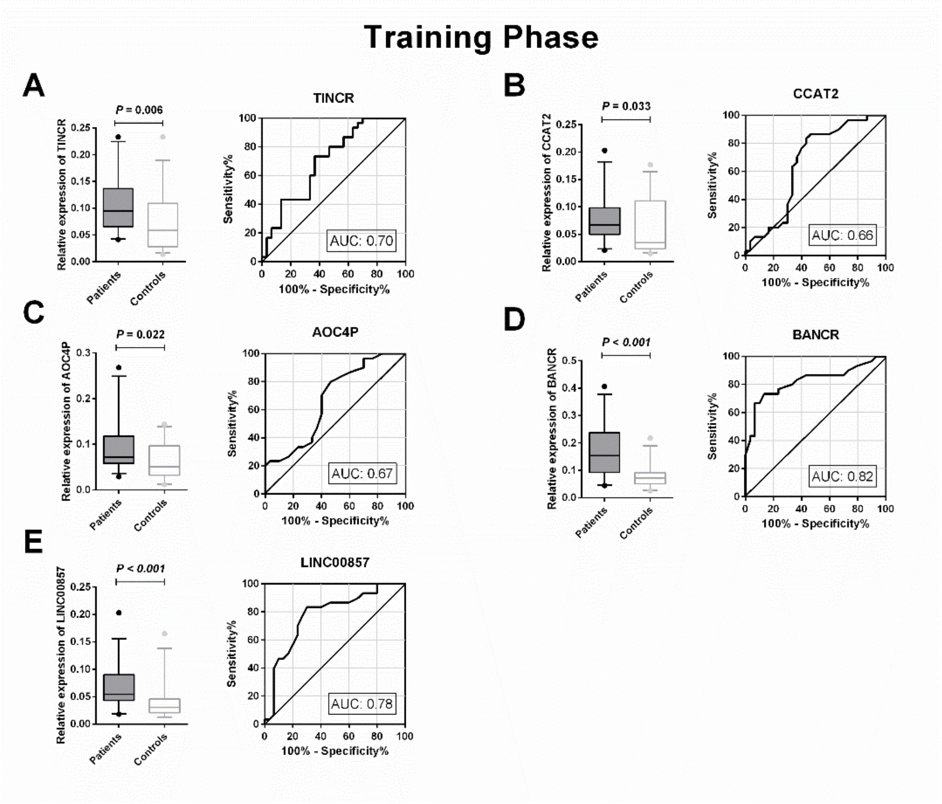

Figure S3. ROC plot for the diagnostic performance of CEA (A), CA125 (B), CA19-9 (C) and CA72-4 (D) in the training phase.

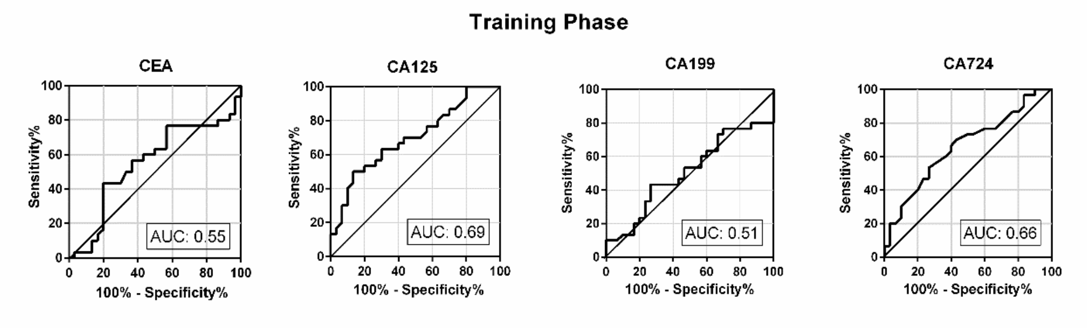

Figure S4. Box and whisker and ROC plots for individual lncRNAs *TINCR* (A), *CCAT2* (B), AOC4P (C), *BANCR* (D) and *LINC00857* (E) in the validation phase. The horizontal line in the middle of each box indicates the median, whereas the top and bottom borders of the box mark the 95th and 5th percentiles, respectively. The whiskers above and below the box extend to the most extreme point no longer than 1.5 times the interquartile range from the box. The Mann-Whitney U test was used.

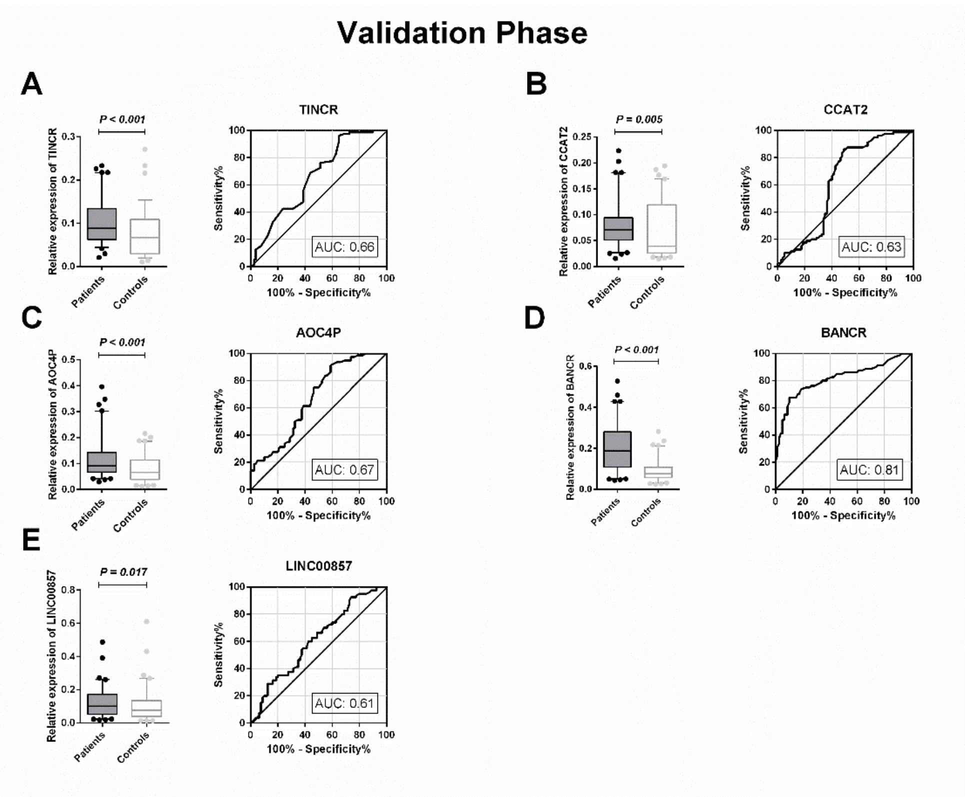

Figure S5. ROC plot for the diagnostic performance of CEA (A), CA125 (B), CA19-9 (C) and CA72-4 (D) in the validation phase.

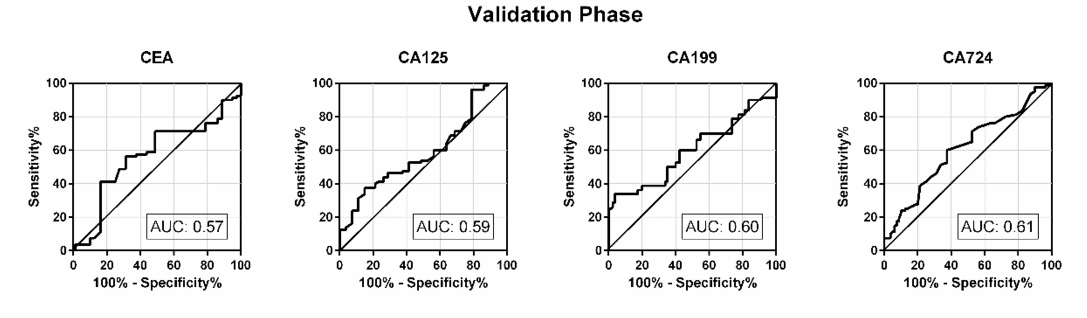

Table S1. qRT-PCR primers

Genes	Forward	Reverse
HOTAIR	5'-GGTAGAAAAAGCAACCACGAAGC-3'	5'-ACATAAACCTCTGTCTGTGAGTGCC-3'
ANRIL	5'-TTGTGAAGCCCAAGTACTGC-3'	5'-TTCACTGTGGAGACGTTGGT-3'
HOXA-AS2	5'-CCCGTAGGAAGAACCGATGA-3'	5'-TTTAGGCCTTCGCAGACAGC-3'
ENST00000601909	5'-CTGTGGAAGTGCTCATCATAGT-3'	5'-CAGCTGTACCACCTACCTTTAC-3'
TINCR	5'-TGTGGCCCAAACTCAGGGATACAT-3'	5'-AGATGACAGTGGCTGGAGTTGTCA-3'
ENST00000509463	5'-CAGTTGGCTCCTGCTAATGT-3'	5'-CTCTTGGTGGAAGCCCTATTC-3'
SPRY4-IT1	5'-GCTGAGCTGGTGGTTGAAAGGAATC-3'	5'-GCTTGGCCCACGATGACTTGG-3'
CCAT2	5'-CCCTGGTCAAATTGCTTAACCT-3'	5'-TTATTCGTCCCTCTGTTTTATGGAT-3'
MIR4435-2HG	5'-CCTGCTCTGCTCCTCTAAATAC-3'	5'-CCGAGGTGGTGATGTTTGA-3'
AOC4P	5'-AAAGGAGGTGAGAGGGAATGT-3'	5'-GCTGGGCACTGGGAGATAC-3'
HOXA10-AS	5'-AGGAAGGGAAGACAGGGAAA -3'	5'-CCAGAACTGGTCGGTGATTTAG-3'
PVT1	5'-TCAAGATGGCTGTGCCTGTC-3'	5'-TTCCACCAGCGTTATTCCCC-3'
BANCR	5'-AGCTTGTATCACCATCTCCATATCA-3'	5'-ACCATAAAAATTATCTGGTCCCTGT-3'
LINC00857	5'-CCCCTGCTTCATTGTTTCCC-3'	5'-AGCTTGTCCTTCTTGGGTACT-3'
LINC00974	5'- TCTAACGTGCCTGGGACCTA-3'	5'- AAATGCCTACCGCCAGTTCA-3'
HULC	5'-ATCTGCAAGCCAGGAAGAGTC-3'	5'-CTTGCTTGATGCTTTGGTCTGT-3'
MALAT1	5'-AACGCAGACGAAAATGGAAAGA-3'	5'-CCTTCTAACTTCTGCACCACCAGA-3'
GAPDH	5'-CATGTTCGTCATGGGTGTGAACCA-3'	5'-AGTGATGGCATGGACTGTGGTCAT-3'
β-actin	5'-TCCTCTCCCAAGTCCACACA-3'	5'-GCACGAAGGCTCATCATTCA-3'

Table S2. Differentially expressed lncRNAs ($P < 0.05$; fold change > 2) collectively selected from two sets of microarray profiling

Dysregulation	Seqname	Gene Symbol	Source	Chrom	Synonyms
up	NR_003716	HOTAIR	RefSeq_NR	Chr12	HOXAS; HOXC-AS4; HOXC11-AS1; NCRNA00072
up	NR_003529	CDKN2B-AS	RefSeq_NR	Chr9	ANRIL; p15AS; PCAT12; CDKN2BAS; CDKN2B-AS1; NCRNA00089
up	NR_122069	HOXA-AS2	RefSeq_NR	Chr7	HOXA3as
up	ENST00000601909	RP11-98D18.17	Ensembl	Chr1	-
up	NR_027064	TINCR	RefSeq_NR	Chr19	PLAC2; LINC00036; NCRNA00036; onco-lncRNA-16
up	ENST00000509463	RP11-597D13.9	Ensembl	Chr4	-
up	NR_002196	H19	RefSeq_NR	Chr11	ASM; BWS; WT2; ASM1; D11S813E; LINC00008; NCRNA00008
up	NR_131221	SPRY4-IT1	RefSeq_NR	Chr5	SPRIGHTLY
up	NR_109834	CCAT2	RefSeq_NR	Chr8	NCCP1; LINC00873
up	ENST00000432818	MIR4435-2HG	Ensembl	Chr2	AGD2; MIR4435-1HG; LINC00978; AK001796
up	NR_002773	AOC4P	RefSeq_NR	Chr17	AOC4; UPAT
up	NR_046609	HOXA10-AS	RefSeq_NR	Chr7	HOXA-AS4
up	NR_003367	PVT1	RefSeq_NR	Chr8	LINC00079; NCRNA00079; onco-lncRNA-100
up	NR_047671	BANCR	RefSeq_NR	Chr9	LINC00586
up	NR_038464	LINC00857	RefSeq_NR	Chr10	-
up	NR_024204	LINC00152	RefSeq_NR	Chr2	C2orf59; NCRNA00152
up	NR_038442	LINC00974	RefSeq_NR	Chr17	-
up	NR_004855	HULC	RefSeq_NR	Chr6	HCCAT1; LINC00078; NCRNA00078

up	NR_002819	MALAT1	RefSeq_NR	Chr11	HCN; NEAT2; PRO2853; mascRNA; LINC00047; NCRNA00047
up	NR_015379	UCA1	RefSeq_NR	Chr19	CUDR; UCAT1; LINC00178; NCRNA00178; onco-lncRNA-36
down	NR_103844	NPTN-IT1	RefSeq_NR	Chr15	lncRNA-LET
down	ENST00000579431	GATA6-AS1	Ensembl	Chr18	locus5689
down	ENST00000609924	RP13-270P17.3	Ensembl	Chr18	-
down	NR_038263	SOCS2-AS1	RefSeq_NR	Chr12	-
down	NR_038170	ABHD16A	RefSeq_NR	Chr6	BAT5; NG26; PP199; D6S82E
down	NR_034105	CRNDE	RefSeq_NR	Chr16	CRNDEP; PNAS-108; lincIRX5; LINC00180; NCRNA00180
down	ENST00000565145	RP11-554A11.9	Ensembl	Chr11	-
down	ENST00000440403	AP000320.6	Ensembl	Chr21	-
down	NR_002766	MEG3	RefSeq_NR	Chr14	GTL2; FP504; prebp1; PRO0518; PRO2160; LINC00023; NCRNA00023;
down	NR_015440	LINC00982	RefSeq_NR	Chr1	-
down	ENST00000590318	RP11-845C23.3	Ensembl	Chr18	-
down	NR_002578	GAS5	RefSeq_NR	Chr1	SNHG2; NCRNA00030
down	ENST00000609497	RP5-1159O4.1	Ensembl	Chr7	-
down	ENST00000447911	SNHG14	Ensembl	Chr15	115HG; LNCAT; UBE3AATS; UBE3A-AS1; NCRNA00214; UBE3A-AS;
down	ENST00000540066	TOPORS-AS1	Ensembl	Chr9	-
down	NR_038883	LINC00649	RefSeq_NR	Chr21	-
down	NR_038285	MFI2-AS1	RefSeq_NR	Chr3	-
down	ENST00000587499	JRK	Ensembl	Chr8	JH8, jerky
down	NR_103734	LINC00342	RefSeq_NR	Chr2	NCRNA00342
down	NR_026761	LINC00467	RefSeq_NR	Chr1	C1orf97
down	NR_038239	LINC00658	RefSeq_NR	Chr20	-
down	NR_033957	LINC00842	RefSeq_NR	Chr10	-
down	ENST00000427857	FAM41C	Ensembl	Chr1	-
down	NR_002323	TUG1	RefSeq_NR	Chr22	TI-227H; LINC00080; NCRNA00080
down	NR_033925	FENDRR	RefSeq_NR	Chr16	FOXF1-AS1; lincFOXF1; TCONS_00024240; onco-lncRNA-21
down	ENST00000562471	RP11-7K24.3	Ensembl	Chr6	-
down	NR_119376	FER1L4	RefSeq_NR	Chr20	C20orf124
down	ENST00000553909	RP11-903H12.5	Ensembl	Chr14	-
down	NR_126369	GACAT1	RefSeq_NR	Chr2	LINC00876
down	NR_004387	SCARNA10	RefSeq_NR	Chr12	U85
down	NR_130107	GHET1	RefSeq_NR	Chr7	lncRNA-GHET1
down	NR_130915	LCAL1	RefSeq_NR	Chr6	onco-lncRNA-27
down	NR_027790	MIR99AHG	RefSeq_NR	Chr21	MONC; C21orf34; C21orf35; FLJ38295; LINC00478
down	ENST00000527061	RP11-867G23.10	Ensembl	Chr11	-
down	ENST00000514752	SLC7A11-AS1	Ensembl	Chr11	-
down	NR_034108	TRAF3IP2-AS1	RefSeq_NR	Chr6	C6UAS; C6orf3; NCRNA00248; TRAF3IP2-AS2
down	ENST00000497872	AL928768.3	Ensembl	Chr14	-
down	ENST00000538616	CD27-AS1	Ensembl	Chr2	-

Table S3. Diagnostic performance of individual biomarkers in the training phase

	Cutoff	Sensitivity (95% CI)	Specificity (95% CI)	AUC (95% CI)	True positive	True negative	False positive	False negative
TINCR	7.2×10^{-2}	0.73 (0.54-0.88)	0.63 (0.44-0.80)	0.70 (0.57-0.84)	22	19	11	8
CCAT2	4.1×10^{-2}	0.83 (0.65-0.94)	0.57 (0.37-0.75)	0.66 (0.52-0.80)	25	17	13	5
AOC4P	5.5×10^{-2}	0.80 (0.61-0.92)	0.53 (0.34-0.72)	0.67 (0.53-0.81)	24	16	14	6
BANCR	11.0×10^{-2}	0.73 (0.54-0.88)	0.87 (0.69-0.96)	0.82 (0.71-0.93)	22	26	4	8
LINC00857	4.0×10^{-2}	0.83 (0.65-0.94)	0.70 (0.51-0.85)	0.78 (0.65-0.90)	25	21	9	5
CEA	5 µg/L	0.13 (0.04-0.31)	0.97 (0.83-100)	0.55 (0.40-0.70)	4	29	1	26
CA125	35 U/ml	0.30 (0.15-0.49)	0.93 (0.78-0.99)	0.69 (0.56-0.83)	9	28	2	21
CA19-9	37 U/ml	0.17 (0.06-0.35)	1.00 (0.48-1.00)	0.51 (0.36-0.66)	5	30	0	25
CA72-4	10 U/ml	0.10 (0.02-0.27)	0.97 (0.83-1.00)	0.66 (0.52-0.80)	3	29	1	27

Table S4. Diagnostic performance of individual biomarkers in the validation phase

	Cutoff[§]	Sensitivity (95% CI)	Specificity (95% CI)	AUC (95% CI)	True positive	True negative	False positive	False negative
TINCR	7.2×10^{-2}	0.69 (0.57-0.79)	0.56 (0.45-0.67)	0.66 (0.57-0.74)	55	45	35	25
CCAT2	4.1×10^{-2}	0.85 (0.75-0.92)	0.51 (0.40-0.63)	0.63 (0.54-0.72)	68	41	39	12
AOC4P	5.5×10^{-2}	0.86 (0.77-0.93)	0.41 (0.30-0.53)	0.67 (0.59-0.75)	69	33	47	11
BANCR	11.0×10^{-2}	0.75 (0.64-0.84)	0.78 (0.67-0.86)	0.81 (0.75-0.88)	60	62	18	20
LINC00857	4.0×10^{-2}	0.93 (0.84-0.97)	0.26 (0.17-0.37)	0.61 (0.52-0.70)	74	21	59	6
CEA	5 µg/L	0.15 (0.08-0.25)	0.89 (0.80-0.95	0.57 (0.48-0.66)	12	71	9	68
CA125	35 U/ml	0.24 (0.15-0.35)	0.93 (0.84-0.97)	0.59 (0.50-0.68)	19	74	6	61
CA199	37 U/ml	0.29 (0.19-0.40)	0.96 (0.89-1.00)	0.60 (0.51-0.69)	23	77	3	57
CA724	10 U/ml	0.11 (0.05-0.20)	0.94 (0.86-0.98)	0.61 (0.52-0.70)	9	75	5	71

[§] Cut-off values identified in the training phase are used