

Supplemental tables

Table S1. The clinical characteristics of the participants in each cohort.

Cohort		Exploration cohort (N)		External validation cohort (N)		Prediction cohort (N)		
		GC (518)	HD (460)	GC (124)	HD (103)	GC (112)	PL (73)	HD (100)
Age, year		54.6±11.0	41.9±13.9	61.3±12.9	48.3±12.5	58.4±13.8	55.9±10.6	53.9±13.4
Gender	Male	339	192	78	62	64	37	51
	Female	179	268	46	41	48	36	49
CEA, ng/ml	Positive	85	12	19	2	13	5	3
	Negative	433	448	105	101	99	68	97
CA19-9, U/ml	Positive	78	6	18	0	15	3	1
	Negative	440	454	106	103	97	70	99
CA72-4, U/ml	Positive	93	55	20	11	20	2	7
	Negative	425	405	104	92	92	71	93
pathological TNM stage	I	46		29		37		
	II	162		36		16		
	III	277		50		47		
	IV	33		9		12		
Maximum diameter, cm	0~1	9		7		8		
	1~5	380		83		77		
	5~10	122		29		26		
	>10	7		5		1		
Tumor location	cardia/fundus	134		50		3		
	gastric body	196		25		48		
	gastric antrum	171		41		42		
	gastric angle	17		7		10		
	whole stomach	0		1		1		
	multiple sizes	0		0		8		
Tumor differentiation	high	5		6		1		
	moderately	43		33		14		
	low	288		83		95		
	signet-ring cell carcinoma	182		2		2		
HER2	positive	12		17		4		
	negative	152		98		93		
	unknown	354		9		15		
Vascular cancer emboli	yes	193		38		42		
	no	269		83		63		
Neural invasion	yes	145		50		55		
	no	318		71		50		

Smoking status	Ever/current	181	0	31	16
	never	337	124	81	56
Drinking status	Ever/current	112	0	19	9
	never	406	124	93	63
Family tumor history	no	436	123	90	17
	yes	82	1	22	55
helicobacter pylori infection	yes	302	20	24	
	no	39	66	5	
	unknown	177	38	83	
BMI, kg/m2		21.9±3.1	20.7±3.3	22.3±3.2	22.7±3.0

Abbreviations: GC, gastric cancer; HD, healthy donor; PL, precancerous lesion.

Table S2. Differential expression of candidate biomarkers in serum exosomes of GC patients and healthy donors in the verification phase.

Biomarker	Mann-Whitney U (GC vs HD)	P value ^a
CTC-327F10.4	147	0.002
CTD-2114J12.1	59	<0.001
KB-1517D11.2	128.5	<0.001
RP11-308D16.1	184.5	0.016
RP11-3B12.5	90	<0.001
RP11-443C10.1	24	<0.001
RP11-556O9.4	64.5	<0.001
RPS14P7	136	0.001
RPS3AP23	163	0.005
OR4F2P	137	<0.001
RPL31P58	147.5	0.002
FYTTD1P1	173	0.009
RP11-417E7.1	123	<0.001
ARL2BPP4	187	0.019
RP11-164H5.1	207	0.049
RNASEH1P1	178	0.012
RP11-318K15.2	147.5	0.002
RP11-589N15.2	129	<0.001
RP11-49I11.3	197	0.031
AIMP1P2	186	0.018
RP11-559M23.1	123	<0.001
CTD-2339L15.3	123	<0.001
IGLVIVOR22-1	94	<0.001
DGCR9	87	<0.001
RP3-395M20.7	94	<0.001
LINC00567	97	<0.001
GK-IT1	78	<0.001
has_circ_0000268	105	<0.001
has_circ_0033144	27	<0.001
has_circ_0047880	96	<0.001
has_circ_0059958	110	<0.001

^a Mann-Whitney test was used for comparing two groups and the significance threshold was a one-tailed P < 0.05.

Abbreviation: GC, gastric cancer; HD, healthy donor.

Table S3. Differential expression of potential biomarkers in GC tissues in comparison with paired normal tissues from the same patients with GC in the verification phase.

Biomarker	t statistics	P value ^a
CTC-327F10.4	3.368	0.001
CTD-2114J12.1	1.872	0.037
KB-1517D11.2	3.230	0.002
RP11-308D16.1	3.324	0.002
RP11-3B12.5	3.063	0.003
RP11-443C10.1	1.760	0.046
RP11-556O9.4	2.759	0.006
RPS14P7	3.119	0.002
RPS3AP23	2.310	0.015
OR4F2P	2.900	0.004
RPL31P58	2.024	0.027
FYTTD1P1	2.086	0.024
RP11-417E7.1	3.306	0.002
ARL2BPP4	3.054	0.003
RP11-164H5.1	2.684	0.007
RNASEH1P1	2.287	0.016
RP11-318K15.2	5.186	<0.001
RP11-589N15.2	3.374	0.001
RP11-49I11.3	2.199	0.019
AIMP1P2	4.134	<0.001
RP11-559M23.1	2.089	0.024
CTD-2339L15.3	2.442	0.011
IGLVIVOR22-1	2.493	0.010
DGCR9	2.298	0.016
RP3-395M20.7	2.223	0.018
LINC00567	2.318	0.015
GK-IT1	2.145	0.021
has_circ_0000268	1.740	0.048
has_circ_0033144	2.796	0.005
has_circ_0047880	6.884	<0.001
has_circ_0059958	4.326	<0.001

^a Paired T test was used for comparing two groups and the significance threshold was a one-tailed P < 0.05.

Abbreviation: GC, gastric cancer.

Table S4. Logistic regression analyses of single non-coding RNA adjusted by age and gender in the training cohort.

Biomarker	OR	95% CI	P value
CTC-327F10.4	53.65	22.90~125.68	<0.001
CTD-2114J12.1	19.80	9.70~40.42	<0.001
KB-1517D11.2	30.66	14.80~63.49	<0.001
RP11-308D16.1	5.90	3.74~9.31	<0.001
RP11-3B12.5	8.67	5.04~14.93	<0.001
RP11-443C10.1	236.69	79.13~707.99	<0.001
RP11-556O9.4	1.91	1.55~2.35	<0.001
RPS14P7	7.96	5.03~12.60	<0.001
RPS3AP23	1.67	1.36~2.05	<0.001
OR4F2P	141.85	54.47~369.40	<0.001
RPL31P58	3.52	2.54~4.87	<0.001
FYTTD1P1	19.79	10.74~36.45	<0.001
RP11-417E7.1	2.43	1.82~3.23	<0.001
ARL2BPP4	10.72	6.47~17.77	<0.001
RP11-164H5.1	29.95	14.97~59.93	<0.001
RNASEH1P1	0.99	0.98~1.01	0.539
RP11-318K15.2	155.35	54.89~439.70	<0.001
RP11-589N15.2	2.98	1.98~4.49	<0.001
RP11-491I1.3	749686.10	43175.40~13017348.00	<0.001
AIMP1P2	10.41	6.43~16.83	<0.001
RP11-559M23.1	319.12	102.59~992.65	<0.001
CTD-2339L15.3	233.14	83.65~649.75	<0.001
IGLVIVOR22-1	20.10	10.13~39.88	<0.001
DGCR9	50.73	24.05~107.03	<0.001
RP3-395M20.7	65.37	26.43~161.69	<0.001
LINC00567	106.46	34.42~329.26	<0.001
GK-IT1	4.32	2.95~6.32	<0.001
has_circ_0000268	7.26	4.51~11.69	<0.001
has_circ_0033144	3.68	2.66~5.09	<0.001
has_circ_0047880	2.07	1.69~2.53	<0.001
has_circ_0059958	2.12	1.58~2.85	<0.001

Only the results of single non-coding RNA in the multivariate model were summarized in this table.

Abbreviation: OR, odds ratio; CI, confidence interval.

Table S5. Performance of CTD.2339L15.3, LINC00567, RP11.443C10.1 and DGCR9 in differentiating patients with gastric cancer from healthy donors.

Cohort	Biomarker	cutoff	accuracy (95%CI)	sensitivity (95%CI)	specificity (95%CI)	F1
Training	CTD.2339L15.3	0.398	0.841 (0.811~0.867)	0.812 (0.772~0.852)	0.873 (0.837~0.908)	0.844
Testing	CTD.2339L15.3	0.398	0.853 (0.807~0.892)	0.821 (0.760~0.881)	0.891 (0.838~0.943)	0.856
Training	LINC00567	0.308	0.858 (0.830~0.884)	0.914 (0.886~0.943)	0.796 (0.752~0.840)	0.872
Testing	LINC00567	0.308	0.843 (0.796~0.883)	0.891 (0.842~0.940)	0.788 (0.720~0.857)	0.858
Training	RP11.443C10.1	0.270	0.838 (0.808~0.865)	0.898 (0.867~0.929)	0.771 (0.725~0.817)	0.854
Testing	RP11.443C10.1	0.270	0.829 (0.781~0.871)	0.865 (0.812~0.919)	0.788 (0.720~0.857)	0.844
Training	DGCR9	0.416	0.826 (0.796~0.854)	0.801 (0.760~0.842)	0.854 (0.816~0.893)	0.830
Testing	DGCR9	0.416	0.788 (0.737~0.834)	0.769 (0.703~0.835)	0.810 (0.745~0.876)	0.795

Abbreviation: CI, confidence interval.

Table S6. Analyses of the impact of clinical characteristics on the cd-score in the exploration cohort.

Types	Factors	Univariate analysis			Multivariate analysis			
		OR	95%CI	<i>P</i> value ^a	OR	95%CI	<i>P</i> value ^a	
Gastric cancer	Sex	female	reference			reference		
		male	0.95	0.66~1.36	0.782	0.92	0.58~1.45	0.709
	Age	<=60 y	reference			reference		
		>60 y	0.71	0.48~1.03	0.070	0.75	0.49~1.14	0.176
	CEA	negative	reference			reference		
		positive	1.22	0.76~1.94	0.407	1.38	0.82~2.32	0.227
	CA19-9	negative	reference			reference		
		positive	0.78	0.48~1.27	0.326	0.80	0.47~1.37	0.417
	CA72-4	negative	reference			reference		
		positive	0.92	0.59~1.45	0.731	0.98	0.60~1.61	0.942
	Maximum diameter	<=5 cm	reference			reference		
		>5 cm	0.80	0.53~1.19	0.264	0.84	0.54~1.30	0.432
	pTNM	I	reference			reference		
		II	1.90	0.97~3.70	0.060	2.19	1.04~4.61	0.039
		III	1.57	0.83~2.96	0.167	1.75	0.85~3.63	0.130
		IV	1.01	0.40~2.53	0.981	0.92	0.31~2.70	0.877
	Differentiation	High	reference			reference		
		Moderately	1.30	0.20~8.61	0.782	0.68	0.09~5.40	0.718
		Low	1.68	0.28~10.18	0.575	0.90	0.13~6.51	0.919
		SRCC	1.31	0.21~8.05	0.767	0.71	0.10~5.19	0.732
	Location	cardia/fundus	reference			reference		
		gastric body	2.24	1.43~3.51	<0.001	2.24	1.35~3.72	0.002
		gastric antrum	1.78	1.12~2.82	0.014	2.15	1.29~3.58	0.003
		gastric angle	1.89	0.69~5.21	0.219	2.65	0.90~7.84	0.077
	Vascular invasion	no	reference			reference		
		yes	0.81	0.56~1.18	0.272	1.03	0.57~1.52	0.778
		unknown	1.27	0.71~2.26	0.426	NA		
	Nerve infiltration	no	reference			reference		
		yes	0.96	0.65~1.42	0.827	0.91	0.55~1.48	0.694
		unknown	1.43	0.80~2.55	0.229	NA		
	Smoking history	never	reference			reference		
		ever/current	1.02	0.71~1.46	0.927	1.24	0.75~2.08	0.403
	Drinking history	never	reference			Reference		
		ever/current	0.76	0.50~1.16	0.201	0.66	0.38~1.12	0.122
	HER2	Positive	reference			reference		
		negative	0.49	0.15~1.62	0.244	0.37	0.10~1.42	0.148
		unknown	0.83	0.26~2.66	0.750	0.45	0.12~1.69	0.237
	Family tumor	no	reference			reference		

Healthy donors	history	yes	0.70	0.44~1.13	0.150	0.74	0.43~1.24	0.251
		<18.5		reference			Reference	
	BMI	18.5-23.9	0.95	0.58~1.55	0.824	0.92	0.54~1.56	0.756
		>24	0.73	0.41~1.29	0.275	0.65	0.35~1.20	0.172
	Lauren classification	Intestinal		reference			reference	
		diffuse	1.11	0.60~2.05	0.751	0.91	0.46~1.83	0.795
		mixed	2.33	0.95~5.74	0.065	1.72	0.66~4.54	0.270
		unknown	2.32	1.34~4.03	0.003	1.82	0.98~3.39	0.058
	helicobacter pylori infection	yes		reference			reference	
		no	1.40	0.72~2.74	0.325	1.51	0.72~3.16	0.272
		unknown	1.17	0.81~1.70	0.402	1.28	0.84~1.95	0.248
	Sex	female		reference			reference	
		male	1.11	0.77~1.61	0.571	1.13	0.78~1.65	0.523
	Age	<=60		reference			reference	
		>60	1.37	0.79~2.39	0.263	1.38	0.79~2.44	0.261
	CEA	negative		reference			reference	
		positive	1.00	0.32~3.15	1	1.13	0.22~5.72	0.882
	CA19-9	negative		reference			reference	
		positive	1.00	0.20~5.01	1	1.13	0.22~5.72	0.882
	CA72-4	negative		reference			reference	
		positive	1.34	0.76~2.36	0.316	1.33	0.75~2.36	0.324

Abbreviations: OR, odds ratio; CI, confidence interval; SRCC, signet-ring cell carcinoma.

Dependent variable: scores of cd-score.

^aWald Test was used to calculate *P* value.

Table S7. Analyses of the impact of clinical characteristics and cd-score on the distribution of transcriptome data in the exploration cohort.

Factors		Univariate analysis			Multivariate analysis		
		OR	95%CI	<i>P</i> value ^a	OR	95%CI	<i>P</i> value ^a
Age	<=60		reference			reference	
	>60	3.01	2.16~4.21	<0.001	2.95	1.71~5.10	<0.001
Gender	female		reference			reference	
	male	2.64	2.04~3.43	<0.001	1.96	1.27~3.03	0.002
CEA	negative		reference			reference	
	positive	7.33	3.95~13.61	<0.001	10.60	3.36~30.15	<0.001
CA19-9	negative		reference			reference	
	positive	13.41	5.79~31.09	<0.001	10.60	3.36~30.15	<0.001
CA72-4	negative		reference			reference	
	positive	1.61	1.12~2.31	0.009	1.29	0.70~2.38	0.419
cd-score	negative		reference			reference	
	positive	76.08	50.28~115.12	<0.001	69.28	44.52~107.81	<0.001

Abbreviations: OR, odds ratio; CI, confidence interval.

Dependent variable: sample status.

^aWald Test was used to calculate *P* value.

Table S8. Analyses of the impact of clinical characteristics on the cd-score in the external validation cohort.

Types	Factors		Univariate analysis			Multivariate analysis		
			OR	95%CI	<i>P</i> value ^a	OR	95%CI	<i>P</i> value ^a
Gastric cancer ^b	Sex	female		reference			reference	
		male	1.15	0.55~2.38	0.710	1.51	0.50~4.57	0.467
	Age	<=60 y		reference			reference	
		>60 y	1.51	0.73~3.12	0.270	1.36	0.40~4.68	0.626
	CEA	negative		reference			reference	
		positive	0.88	0.33~2.35	0.803	1.97	0.41~9.59	0.400
	CA19-9	negative		reference			reference	
		positive	1.69	0.61~4.71	0.311	0.85	0.17~4.30	0.846
	CA72-4	negative		reference			reference	
		positive	1.62	0.61~4.29	0.332	1.17	0.31~4.39	0.815
	Maximum diameter	<=5 cm		reference			reference	
		>5 cm	1.18	0.53~2.59	0.687	1.59	0.49~5.18	0.440
	pTNM	I		reference			reference	
		II	1.27	0.47~3.40	0.638	1.19	0.25~5.69	0.831
		III	1.53	0.61~3.87	0.364	0.82	0.18~3.83	0.800
		IV	4.96	0.87~28.15	0.071	7.63	0.50~116.26	0.143
	Location	cardia/fundus		reference			reference	
		gastric body	1.08	0.41~2.83	0.870	1.33	0.31~5.81	0.702
		gastric antrum	1.50	0.65~3.44	0.339	0.91	0.28~2.88	0.867
		gastric angle	0.88	0.18~4.35	0.876	0.42	0.05~3.74	0.433
	Vascular invasion	no		reference			reference	
		yes	1.63	0.75~3.54	0.218	1.97	0.56~6.91	0.289
	Nerve infiltration	no		reference			reference	
		yes	2.05	0.98~4.28	0.056	2.26	0.71~7.15	0.166
	HER2	Positive		reference			Reference	
		negative	2.71	0.89~8.28	0.080	11.39	1.85~70.30	0.009
		unknown	3.00	0.56~16.07	0.200	1.85	0.15~23.40	0.634
	BMI	<18.5		reference			reference	
		18.5-23.9	0.69	0.29~1.63	0.396	0.62	0.19~2.03	0.430
		>24	0.56	0.17~1.85	0.346	0.96	0.18~5.10	0.964
	Lauren classification	Intestinal	0.53	0.20~1.42	0.207	0	NA	0.990
		diffuse	1.06	0.43~2.60	0.903	0.39	0.10~1.52	0173
		mixed		reference			reference	
		unknown	2.62	0.24~28.20	0.426	0	NA	0.993
	helicobacter pylori infection	yes		reference			reference	
		no	0.69	0.24~1.94	0.479	0.98	0.24~3.92	0.976
		unknown	0.25	0.08~0.78	0.017	0.06	0.01~0.38	0.003
Healthy	Sex	female		reference			reference	

donors ^c	male	0.49	0.22~1.10	0.085	0.48	0.21~1.08	0.075
Age	<=60	reference		reference			
	>60	1.36	0.44~4.25	0.593	1.53	0.48~4.91	0.470
CEA	negative	reference		reference			
	positive	0.98	0.06~16.11	0.989	1.33	0.07~23.60	0.847
CA72-4	negative	reference		reference			
	positive	1.20	0.34~4.21	0.776	1.15	0.31~4.20	0.836

Abbreviations: OR, odds ratio; CI, confidence interval; SRCC, signet-ring cell carcinoma.

Dependent variable: scores of cd-score.

^aWald Test was used to calculate *P* value.

^bThe numbers of patients with gastric cancer in high differentiation, with smoking history, with drinking history and with family tumor history were too small for analysis.

^cAll heathy donors were negative for CA19-9 in the external validation cohort.

Table S9. Performance of the cd-score in differentiating patients with negative CEA, CA19-9 and CA72-4 from healthy donors.

Cohort	cutoff	Accuracy (95%CI)	sensitivity (95%CI)	specificity (95%CI)	F1
Training cohort	0.473	0.902 (0.874~0.926)	0.899 (0.860~0.938)	0.904 (0.872~0.936)	0.884
Testing cohort	0.473	0.879 (0.830~0.918)	0.853 (0.781~0.924)	0.898 (0.847~0.949)	0.853
External validation cohort	0.473	0.876 (0.818~0.920)	1 (1~1)	0.786 (0.707~0.866)	0.871

Abbreviations: cd-score, score of combined diagnostic model; CI, confidence interval.

Table S10. Performance of cd-score, CEA, CA19-9 and CA72-4 in the prediction cohort.

cohort	Biomarker	cutoff	Accuracy (95%CI)	sensitivity (95%CI)	specificity (95%CI)	F1
GC vs HD (112 vs 100)	cd-score	0.473	0.958 (0.921~0.980)	1 (1~1)	0.910 (0.854~0.966)	0.961
	CEA	5	0.542 (0.473~0.611)	0.357 (0.268~0.446)	0.750 (0.665~0.835)	0.452
	CA19-9	35	0.538 (0.468~0.606)	0.134 (0.071~0.197)	0.990 (0.970~1.000)	0.234
	CA72-4	6.9	0.533 (0.463~0.602)	0.179 (0.108~0.250)	0.930 (0.880~0.980)	0.288
GC vs PL (112 vs 73)	cd-score	0.537	0.962 (0.924~0.985)	0.991 (0.974~1.000)	0.918 (0.855~0.981)	0.969
	CEA	5	0.508 (0.434~0.582)	0.357 (0.268~0.446)	0.740 (0.639~0.840)	0.468
	CA19-9	35	0.459 (0.386~0.534)	0.134 (0.071~0.197)	0.959 (0.913~1.000)	0.231
	CA72-4	6.9	0.492 (0.418~0.566)	0.179 (0.108~0.250)	0.973 (0.935~1.000)	0.299
PL vs HD (73 vs 100)	cd-score	0.462	0.757 (0.686~0.819)	0.603 (0.490~0.715)	0.870 (0.804~0.936)	0.677
	CEA	5	0.543 (0.466~0.619)	0.260 (0.160~0.361)	0.750 (0.665~0.835)	0.325
	CA19-9	35	0.512 (0.512~0.664)	0.041 (0.000~0.087)	0.990 (0.970~1.000)	0.078
	CA72-4	6.9	0.549 (0.472~0.625)	0.027 (0~0.065)	0.930 (0.880~0.980)	0.049

Abbreviations: CI, confidence interval; cd-score, score of combined diagnostic model; GC, gastric cancer; PL, precancerous lesion; HD, healthy donor.