1 **Integrated Molecular Profiling of Human Gastric Cancer Identifies DDR2 as a**

2 **Potential Regulator of Peritoneal Dissemination**

3

4 Junji Kurashige[1,2], Takanori Hasegawa[3], Atsushi Niida[4], Keishi Sugimachi[1], Niantao

5 Deng[5], Kosuke Mima[1,2], Ryutaro Uchi[1], Genta Sawada[1,6], Yusuke Takahashi[1,6],

6 Hidetoshi Eguchi[1], Masashi Inomata[7], Seigo Kitano[7], Takeo Fukagawa[8], Mitsuru

7 Sasako[9], Hiroki Sasaki[10], Shin Sasaki[11], Masaki Mori[6], Kazuyoshi Yanagihara[12], Hideo

8 Baba[2], Satoru Miyano[4], Patrick Tan[5], Koshi Mimori[1]

9

10 **Supplementary Figure Legends**

11 **Supplementary Figure 1**

12 **Establishment of high peritoneal disseminated gastric cell lines.**

13 (a) Four cell lines with high peritoneal disseminated characteristics were isolated from

14 each parental cell using 12 repeating cycles of orthotropic transplantation directly into

15 the gastric wall in nude mice. (b) Mice implanted with parental HSC-58 cells had few, if

16 any, nodules and no apparent accumulation of ascites (left), and those implanted with

17 disseminated 58As9Luc cells formed disseminated nodules with ascites in the

18 peritoneum (right). Red arrows are tumors on the gastric wall, and the white arrows are

19 peritoneal nodules. (c) H&E staining of histological sections from an orthotopic tumor

1   formed by HSC-58 at gastric wall (top), an orthotopic tumor formed by 58As9Luc cells

2   at gastric wall (middle), and peritoneal nodules formed by 58As9Luc cells (bottom).

3   Scale bar, 50 µm. (d) 40 sets of the genes are highly (top 50, 100, 200, and 400)

4   upregulated or downregulated between each of the four pairs of parent and derived cell

5   lines and also between all parents and derived.

6

7   **Supplementary Figure 2**

8   **A flow chart of our EEM analysis**

9   At first, we performed EEM analysis of the 40 gene sets in the gastric tumor expression

10   data set (200 Singapore clinical set) and then obtained 18 significant gene sets by a

11   threshold ($p < 1.0e-6$). Second, we additionally applied EEM analysis to 1,981 gene sets

12   that are registered as the GO and curatedCp entries in MSigDB, namely, 2,021

13   (=40+1,981) gene sets in total. Among these gene sets, only 139 gene sets were selected

14   as significant by a threshold ($p < 1.0e-6$). The significant cell line-derived gene sets

15   were then subject to statistical tests while all significant gene sets were subjected to a

16   clustering analysis.

17

18   **Supplementary Figure 3**

1   **DDR2 had a higher expression in 58As9Luc than in HSC-58 cell line.**
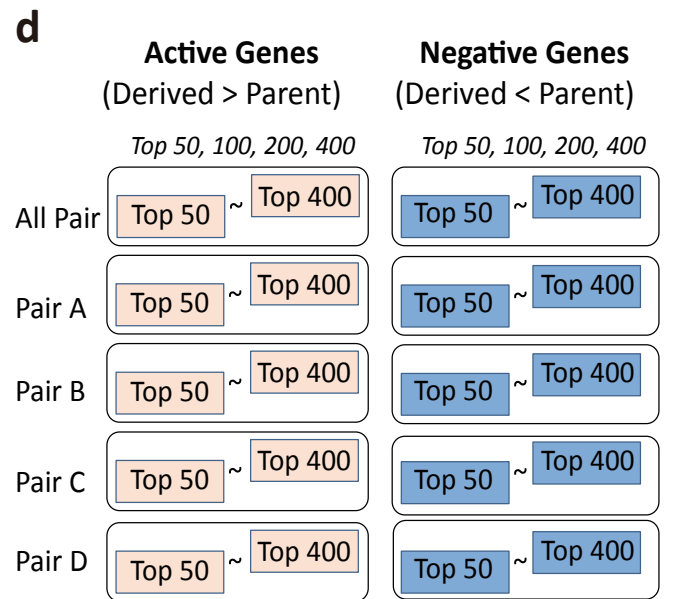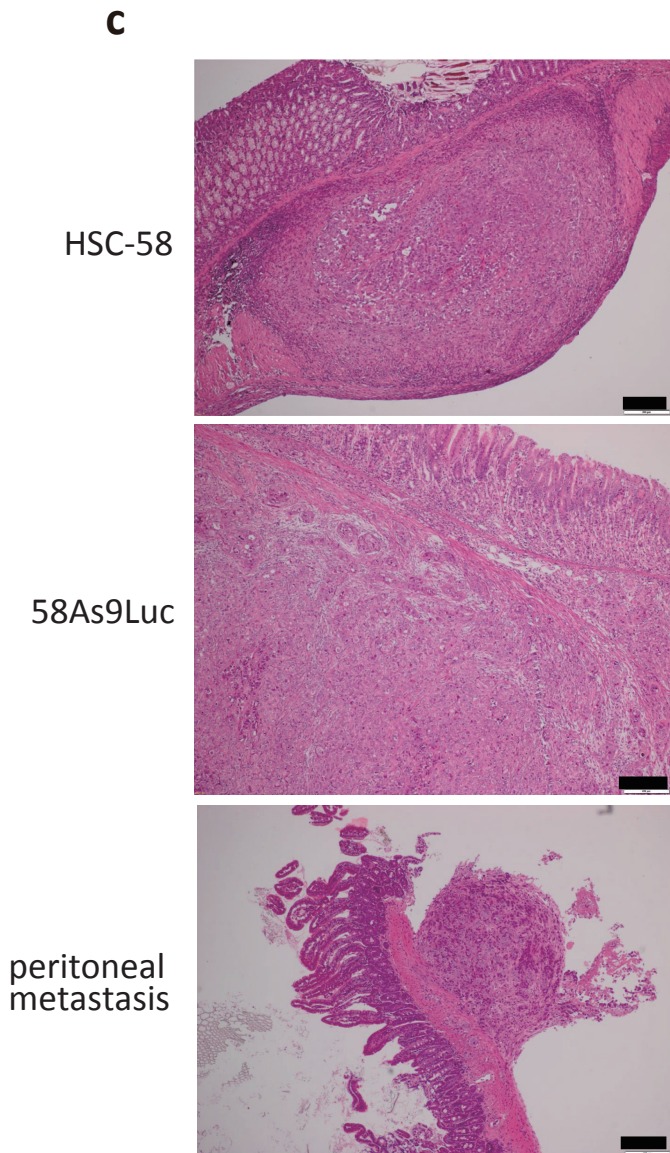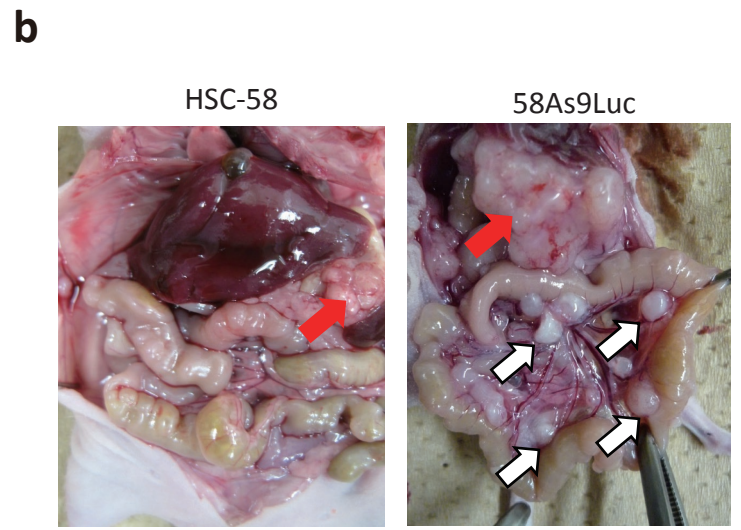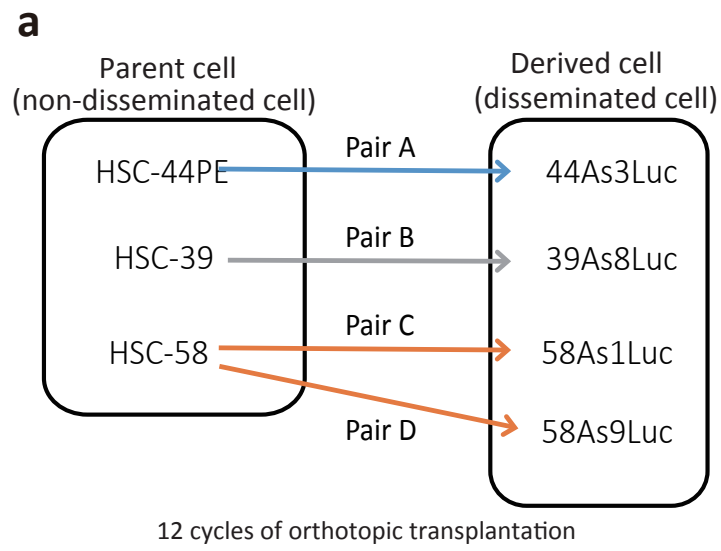
2   (a) The DDR2 expression in HSC-58 and 58As9Luc were confirmed by qRT-PCR, (b)

3   by western blotting, and (c) by immunohistochemistry of the mouse model.
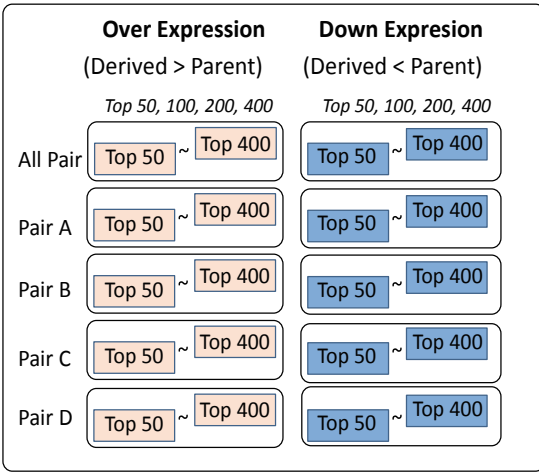
4

5   **Supplementary Figure 4**

6   **Suppression of DDR2 reduced migration and invasion in MKN7 cell line.**

7   (a) The expression of *DDR2* mRNA in 10 gastric cell lines. (b) *DDR2* expression was

8   suppressed in 58As9Luc by two independent shRNA in MKN7. (c) Effects of *DDR2* on

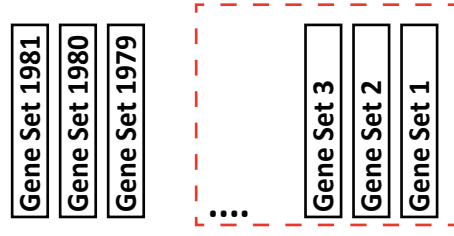9   migration and (d) invasion in MKN7.

a

Parent cell
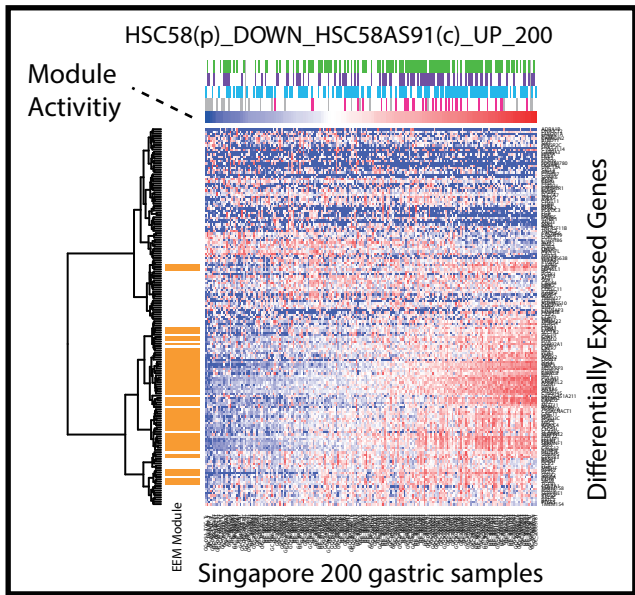(non-disseminated cell)

Derived cell
(disseminated cell)

HSC-44PE — Pair A → 44As3Luc

HSC-39 — Pair B → 39As8Luc

HSC-58 — Pair C → 58As1Luc

HSC-58 — Pair D → 58As9Luc

12 cycles of orthotopic transplantation

b

HSC-58

58As9Luc

c

HSC-58

58As9Luc

peritoneal
metastasis

d

	Active Genes (Derived > Parent)		Negative Genes (Derived < Parent)	
	Top 50, 100, 200, 400		*Top 50, 100, 200, 400*	
All Pair	Top 50	~ Top 400	Top 50	~ Top 400
Pair A	Top 50	~ Top 400	Top 50	~ Top 400
Pair B	Top 50	~ Top 400	Top 50	~ Top 400
Pair C	Top 50	~ Top 400	Top 50	~ Top 400
Pair D	Top 50	~ Top 400	Top 50	~ Top 400

40 Gene Sets from Cell Line

	Over Expression (Derived > Parent)	Down Expression (Derived < Parent)
	Top 50, 100, 200, 400	*Top 50, 100, 200, 400*
All Pair	Top 50 ~ Top 400	Top 50 ~ Top 400
Pair A	Top 50 ~ Top 400	Top 50 ~ Top 400
Pair B	Top 50 ~ Top 400	Top 50 ~ Top 400
Pair C	Top 50 ~ Top 400	Top 50 ~ Top 400
Pair D	Top 50 ~ Top 400	Top 50 ~ Top 400

1,981 Gene Sets (GO and curatedCp)

Singapore Expression

EEM

Gene Set 1981 Gene Set 1980 Gene Set 1979 .... Gene Set 3 Gene Set 2 Gene Set 1

Gene Sets with p < 1.0e-6 (139 gene sets)

-Gene Sets with Coherent Subset and p-value
-Average Expression of Coherent Subset (module activity)

ex) EEM Result

HSC58(p)_DOWN_HSC58AS91(c)_UP_200

Module Activitiy

Differentially Expressed Genes

EEM Module

Singapore 200 gastric samples

Singapore Clinical Data

Gene Sets from Cell Line with p < 1.0e-6 (18 gene sets)

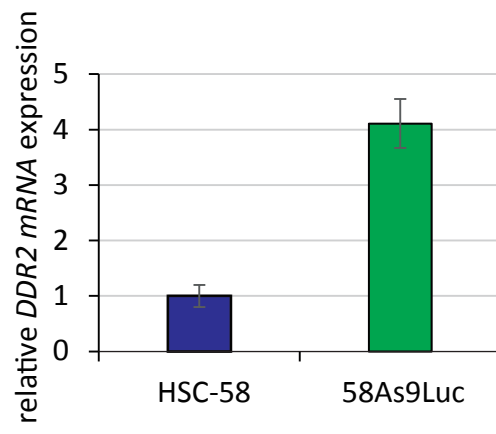**Gene Sets with p < 1.0e-6 (139 gene sets)**

Survival Analysis

Clustering Analysis

Statistical Test (module activity *vs* clinical data)

-Kaplan-Meier method (overall survival)

Student's t-test (peritoneal metatasis)

log-rank *p* = 0.00895

Overall Survival ratio

Years after operation (year)

— Module Activity > 0 n =99
— Module Activity < 0 n=99

Module Activity
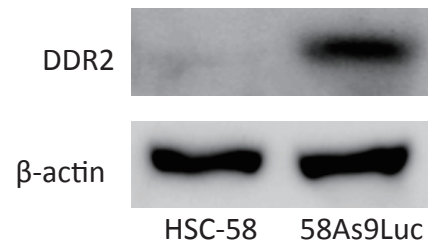
p = 0.00066

peritoneal metastasis (—) n = 141   peritoneal metastasis (+) n = 32
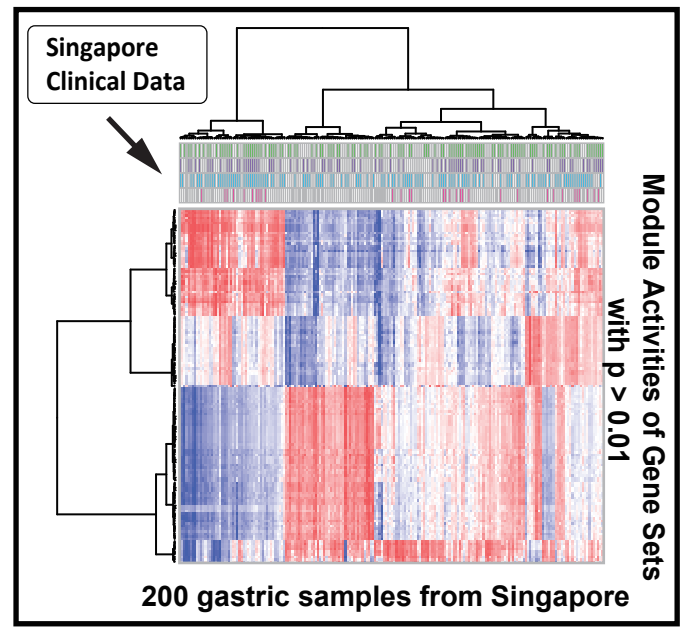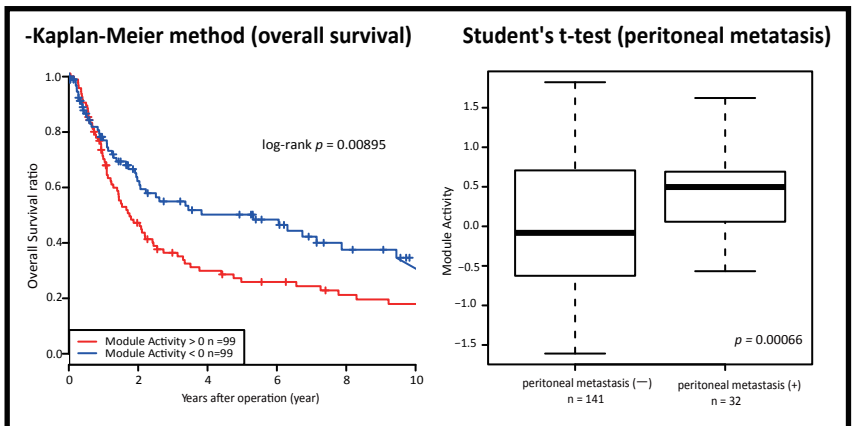
The most significant gene set among 18 Gene Sets

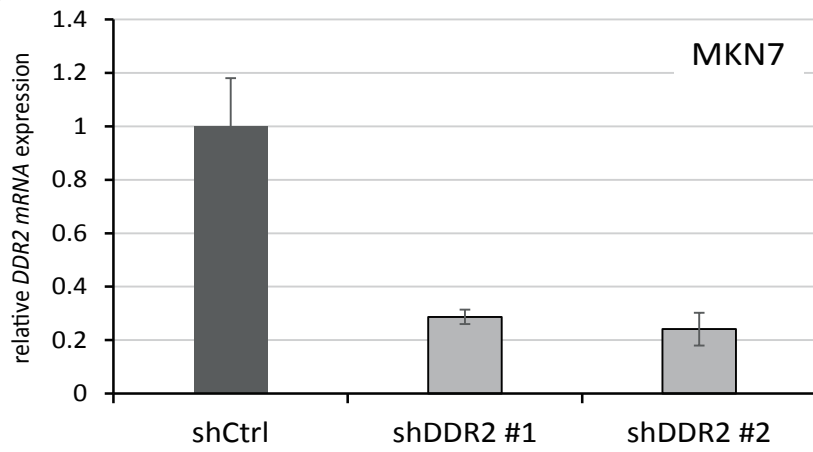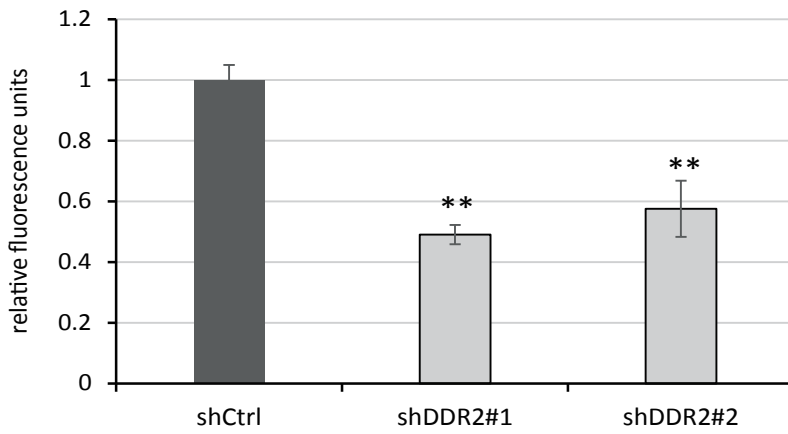**Singapore Clinical Data**

Module Activities of Gene Sets with p > 0.01

200 gastric samples from Singapore

a

relative *DDR2 mRNA* expression

b

DDR2

β-actin

HSC-58 58As9Luc

c

DDR2

HSC-58 58As9Luc