SUPPLEMENTAL INFORMATION

Methods and Materials

CELL CULTURE

MDA-MB-231, a human breast cancer cell line, was obtained from the American Type Culture Collection. The cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) containing penicillin (50 U/ml), streptomycin (50 μg/ml), and glutamine (1 mg/ml), and supplemented with 10% fetal bovine serum, and maintained under a moist 5%-CO_2/95%-air atmosphere and fresh medium exchanged every 2 days. Confluent cells were washed twice with phosphate-buffered saline (PBS, pH 7.4) and harvested by a brief incubation in PBS with a trypsin-EDTA solution (Sigma Aldrich).

COLLECTION OF CONDITIONED CELL-CULTURE MEDIUM (CM)

MDA- MB-231 cells were grown to 80% confluence in DMEM medium supplemented with 10% fetal bovine serum, then washed in serum-free medium three times and incubated in serum-free medium for 18 hours. The CM was centrifuged to remove any cellular debris; and then concentrated through a centrifugal filter (3000 MW cut off, Millipore, USA). The protein concentrated of the sample was determined and then stored at -80 °C until use.

IMMUNOBLOTTING

Protein samples (10 μg or 20 μg from interstitial fluids, 2 μg from sera) were subjected to SDS-PAGE electrophoresis. Proteins were transferred to a polyvinylidene fluoride membrane and immunoblotted with a polyclonal antibody against carboxypeptidase N subunit 2 (Anti-CPN,

1:200, Abcam, USA). The signal was detected by chemiluminescence using the Pierce ECL Western Blotting kit (Thermo Scientific, Rockford, IL, USA).

IMMUNOHISTOCHEMISTRY

Paraffin-embedded tissue (normal or diseased) was cut into 4-μm sections and adhered onto glass slides. The human breast tissue sections already prepared on glass slides were obtained commercially (BC081116a, US BioMax, Rockville, MD). The paraffin in these sections was removed with xylenes, and the samples were rehydrated by changing alcohols with distilled water in step-wise and automated fashion using the Shandon-Lipshaw Varistain (Pittsburgh, PA, USA). The sections were incubated at 90 °C for 10 min in 10 mmol/L citrate buffer, pH 6.0. The anti-CPN polyclonal antibody (Abcam) was applied at a 1:20 dilution in 1% bovine serum albumin. The degree of staining was evaluated by a pathologist, according to the method reported by Zhang et al. *(31)*.

BLOOD SAMPLE COLLECTION FROM HUMAN

Blood samples were collected from healthy women without a history of breast cancer (i.e., non-breast cancer or otherwise "healthy") (36 to 85 years old) and from breast cancer patients (44 to 88 years old) diagnosed with breast cancer at different pathological stages (Table S2), who were enrolled at the University of Texas MD Anderson Cancer Center (MDACC). All participants signed an informed consent form allowing banking of their blood samples for future research. The banking protocol and a protocol describing this research were approved by the MDACC Institutional Review and Privacy Board. Details on patient age and pathologic diagnosis are given in Table S1. From each patient or control individual, 10 mL of whole blood were collected

in plain tubes containing EDTA (Sarstedt, Newton, NC). The specimens were centrifuged at 20 °C for 10 min at 3000 × g, after which the plasma was removed, divided into aliquots, and stored at –80 °C until further use.

ANALYSIS BY MALDI-TOF MS

Each sample (0.5 µl) eluted from the NPS chip was spotted on the MALDI plate and allowed to dry completely, followed by spotting 0.5 µl of the matrix solution (4 g/l of α-cyano-4-hydroxycinnamic acid in a 50% ACN and 0.1% trifluoroacetic acid solution) on top of each sample. After the matrix has crystallized, the plate was subjected to the Applied Biosystems 4700 Proteomics Analyzer (Framingham, MA), operating in positive ion (ion source) and reflective (detector) modes.

ANALYSIS BY HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY-TANDEM MASS SPECTROMETRY (HPLC–MS/MS)

HPLC-MS/MS was processed by the Orbitrap-XL mass spectrometer (Thermo Scientific, Waltham MA) coupled with an Agilent 1200 series HPLC system (Santa Clara, CA). The mobile phases used for LC analysis were (A) 0.1% formic acid in water and (B) 0.1% formic acid in ACN. Flow rate was 0.4 µl/min. MS conditions included a voltage of 2.9 kV for electrospray ionization, an isolation window for MS/MS of 3, relative collision energy of 35%, and a scan strategy of 1 Fourier Transform Mass Spectrometer (FTMS) scan followed by 3 MS/MS product ion scans. Spectra were searched against the SwissProt protein database using Mascot (v 2.3, Matrix Science, London, UK) or Sequest (Thermo Scientific). Search parameters included no

enzyme, mass tolerances of ± 15 ppm for precursors and ± 0.8 Da for fragments, and variable modifications (oxidation of methionine and deamidation of asparagine and/or glutamine).

Table S1. Circulating peptides resulting from the cleavage of the C-terminal arginine by CPN in human plasma and mouse serum.

Peptide name	Peptide sequence	m/z	Protein	Species
C3f_R_{1310}-L_{1319}	RIHWESASLL	1211.65	Complement C3	Human
C3f_R_{1309}-L_{1319}	HRIHWESASLL	1348.70		
C3f_R_{1308}-L_{1319}	THRIHWESASLL	1449.80		
C3f_R_{1305}-L_{1319}	SKITHRIHWESASLL	1777.96		
C3f_R_{1304}-L_{1319}	SSKITHRIHWESASLL	1864.98		
BK_8	RPPGFSPF	904.48	Bradykinin	
C3f_R_{1310}-L_{1319}	RLLWENGNLL	1227.68	Complement C3	Mouse
C3f_R_{1309}-L_{1319}	FR LLWENGNLL	1374.75		
C3f_R_{1308}-L_{1319}	TFR LLWENGNLL	1475.80		
C3f_R_{1307}-L_{1319}	TTFR LLWENGNLL	1576.84		
C3f_R_{1304}-L_{1319}	SSATTFR LLWENGNLL	1821.95		
BK_8	RPPGFSPF	904.48	Bradykinin	

Table S2. Clinical information of the human plasma samples from healthy women and from patients diagnosed with breast cancer at different pathological stages. NOS, not otherwise specified**.**

Woman plasma ID	Stage at Sample Collection	Age at Sample Collection	Note
1	×	36	×
8	×	42	×
4	×	44	×
10	×	50	×
3	×	61	×
5	×	61	×
7	×	64	×
9	×	68	×
6	×	72	×
2	×	85	×
11	I	61	Invasive Ductal Carcinoma, NOS
12	I	50	Invasive Ductal Carcinoma, NOS
13	I	62	Invasive Ductal Carcinoma, NOS
14	I	47	Invasive Lobular Carcinoma, NOS
15	I	88	Invasive Mixed Ductal / Lobular
16	I	45	Invasive Ductal Carcinoma, NOS
17	I	64	Invasive Ductal Carcinoma, NOS
18	I	49	Invasive Ductal Carcinoma, NOS
19	I	65	Invasive Ductal Carcinoma, NOS
20	I	64	Invasive Lobular Carcinoma, NOS
21	I	36	Invasive Ductal Carcinoma, NOS
22	IIA	52	Invasive Ductal Carcinoma, NOS
23	IIA	45	Invasive Ductal Carcinoma, NOS
24	IIA	45	Invasive Ductal Carcinoma, NOS
25	IIA	56	Invasive Ductal Carcinoma, NOS
26	IIA	86	Invasive Lobular Carcinoma, NOS
27	IIA	63	Invasive Ductal Carcinoma, NOS
28	IIB	74	Invasive Mixed Ductal / Lobular
29	IIB	58	Invasive Ductal Carcinoma, NOS
30	IIB	53	Invasive Ductal Carcinoma, NOS
31	IIB	44	Invasive Ductal Carcinoma, NOS
32	IIB	42	Invasive Ductal Carcinoma, NOS
33	IIIA	61	Invasive Ductal Carcinoma, NOS
34	IIIA	67	Invasive Lobular Carcinoma, NOS
35	IIIA	66	Invasive Ductal Carcinoma, NOS
36	IIIB	65	Invasive Ductal Carcinoma, NOS
37	IIIB	64	Invasive Ductal Carcinoma, NOS
38	IIIB	42	Invasive Ductal Carcinoma, NOS
39	IIIC	62	Invasive Lobular Carcinoma, NOS
40	IIIC	42	Invasive Ductal Carcinoma, NOS
41	IIIC	43	Invasive Ductal Carcinoma, NOS
42	IIIC	42	Invasive Lobular Carcinoma, NOS
43	IIIC	41	Invasive Ductal Carcinoma, NOS
44	IIIC	38	Invasive Ductal Carcinoma, NOS
45	IIIC	75	Invasive Ductal Carcinoma, NOS
46	IIIC	78	Invasive Lobular Carcinoma, NOS
47	IV	76	Invasive Ductal Carcinoma, NOS
48	IV	73	Invasive Ductal Carcinoma, NOS
49	IV	48	Invasive Mixed Ductal / Lobular
50	IV	58	Invasive Ductal Carcinoma, NOS
51	IV	65	Invasive Ductal Carcinoma, NOS
52	IV	76	Invasive Mixed Ductal / Lobular
53	IV	48	Invasive Ductal Carcinoma, NOS
54	IV	75	Invasive Lobular Carcinoma, NOS
55	IV	70	Invasive Lobular Carcinoma, NOS
56	IV	72	Invasive Ductal Carcinoma, NOS
57	IV	67	Invasive Lobular Carcinoma, NOS
58	IV	73	Invasive Mixed Ductal / Lobular

Figure S1

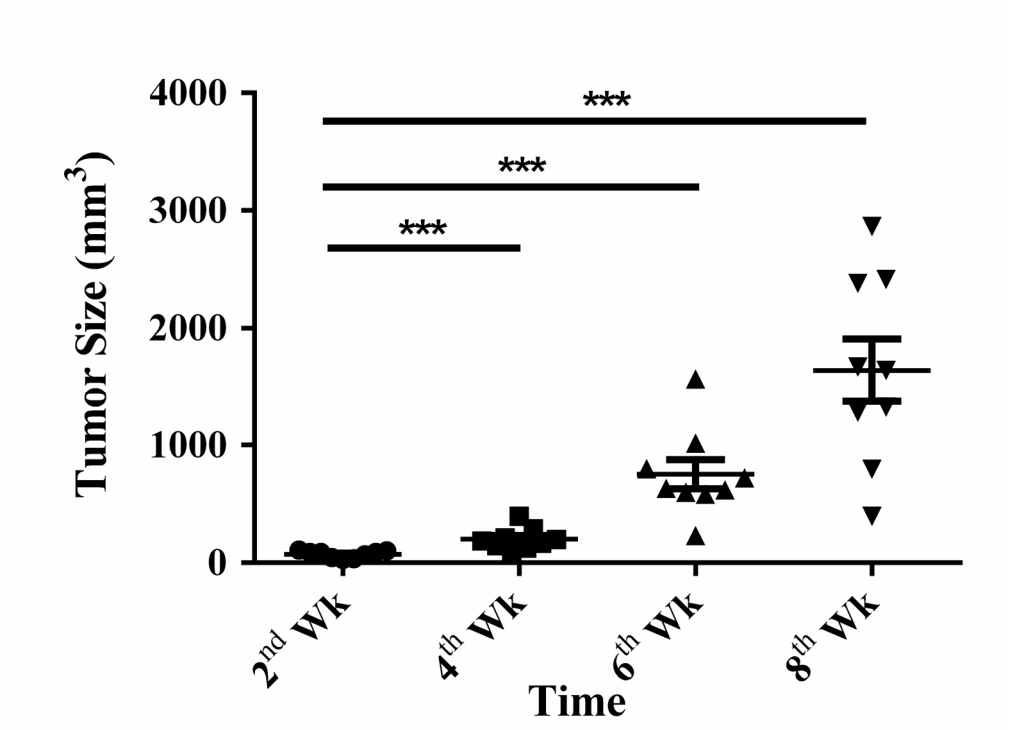

Figure S1. Dot plot for the tumor sizes versus time in breast cancer mouse model. Mouse number: n=8; Student's t-test, ** $p<0.01$, *** $p<0.001$.

Figure S2

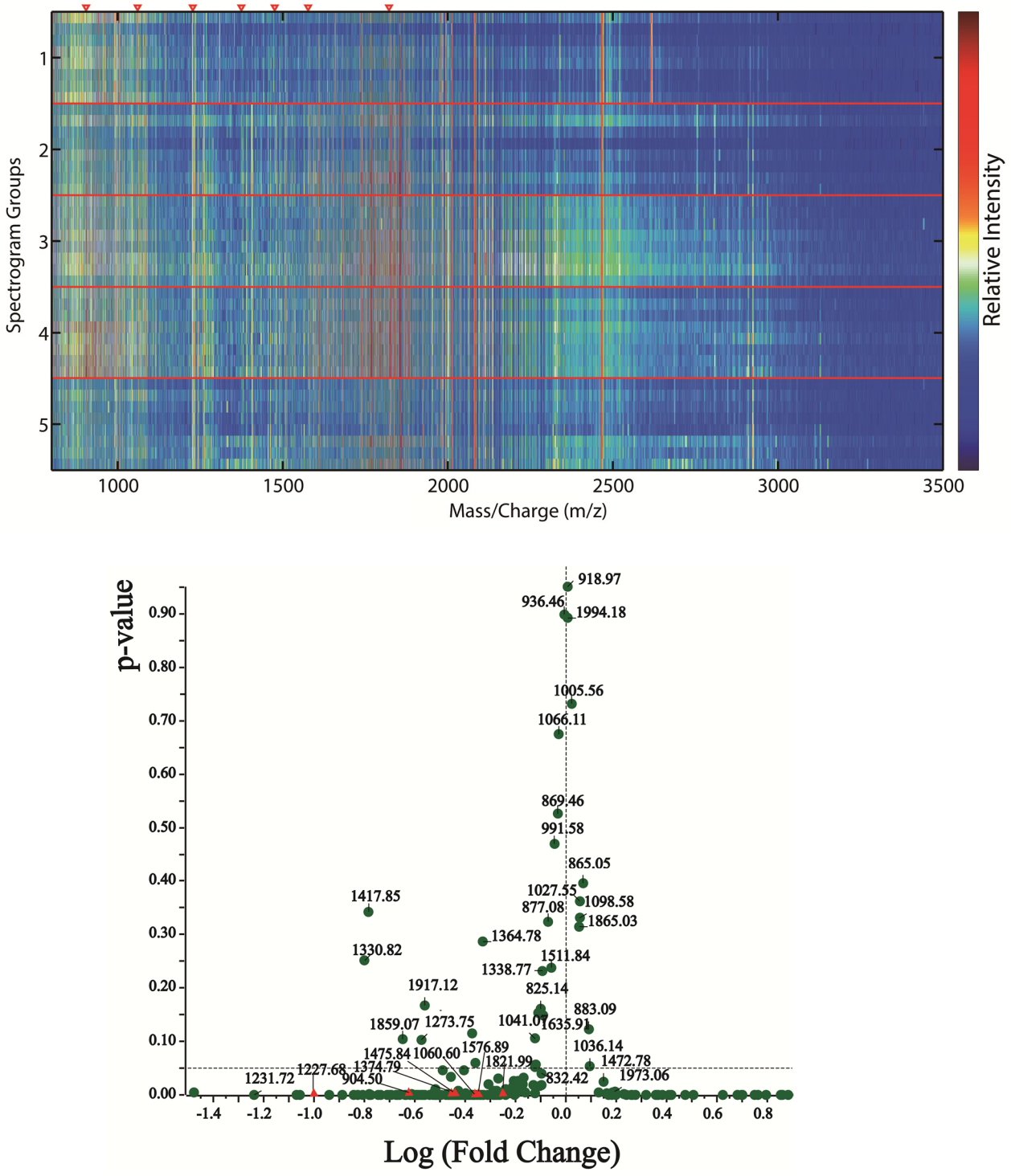

Figure S2. Heat map and dot plot for analysis of the MALDI-TOF MS spectra on circulating peptides from breast cancer mouse model.

A. Heat map. Spectrum group 1-5 indicate the five time points of the blood collection from mice at 0, 2nd, 4th, 6th, 8th week, respectively. n=8/group. **B.** Dot plot (p-value versus log(fold change)).Student's t-test was processed on MALDI-TOF MS spectra collected from mouse sera (Group: 0, 2nd, 4th, 6th, 8th week). Peaks of interests are labeled by red triangles: 904.50, 1060.60, 1227.68, 1374.79, 1475.84, 1576.89 and 1821.99, separately.

Figure S3A

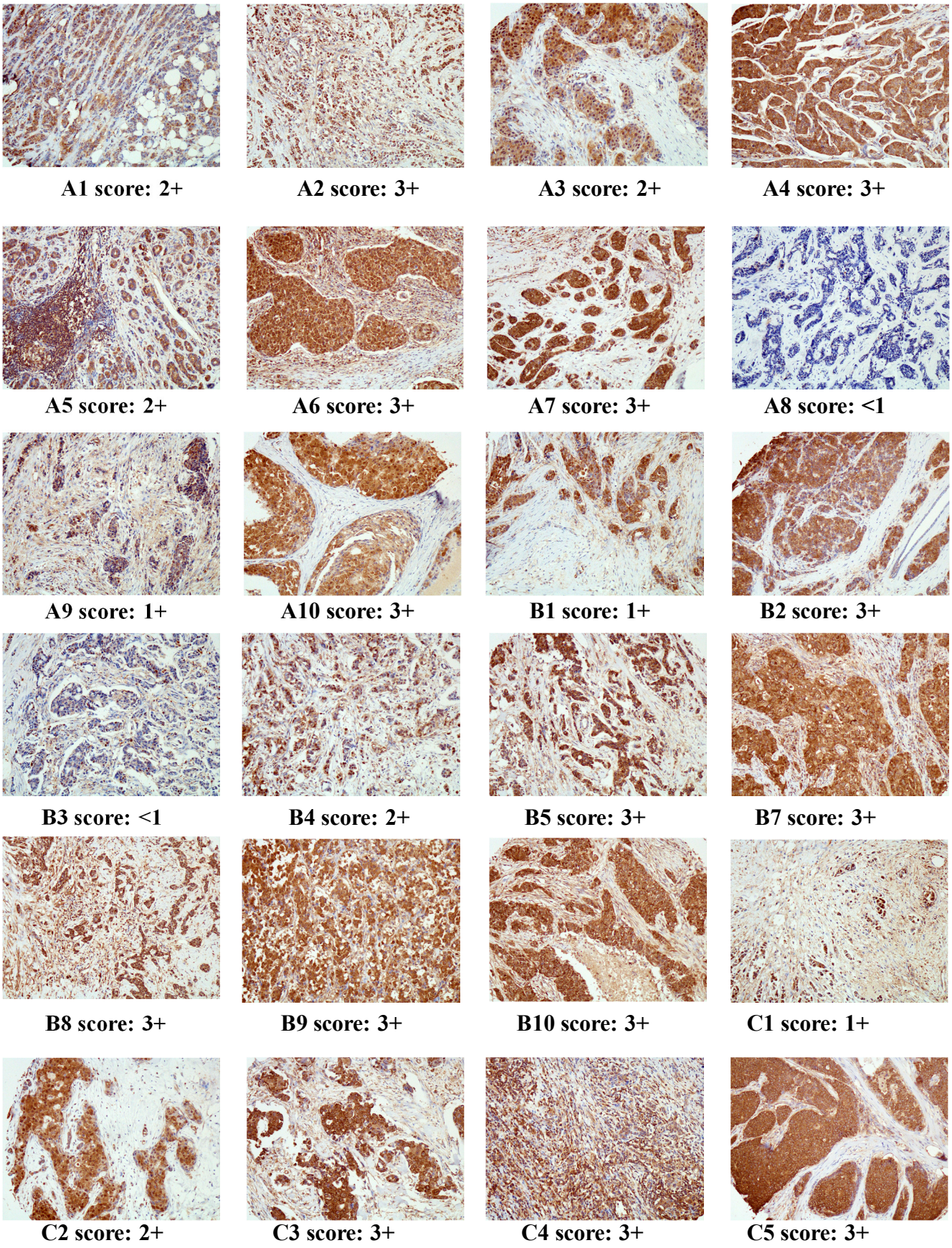
A1 score: 2+
A2 score: 3+
A3 score: 2+
A4 score: 3+
A5 score: 2+
A6 score: 3+
A7 score: 3+
A8 score: <1
A9 score: 1+
A10 score: 3+
B1 score: 1+
B2 score: 3+
B3 score: <1
B4 score: 2+
B5 score: 3+
B7 score: 3+
B8 score: 3+
B9 score: 3+
B10 score: 3+
C1 score: 1+
C2 score: 2+
C3 score: 3+
C4 score: 3+
C5 score: 3+

Figure S3B

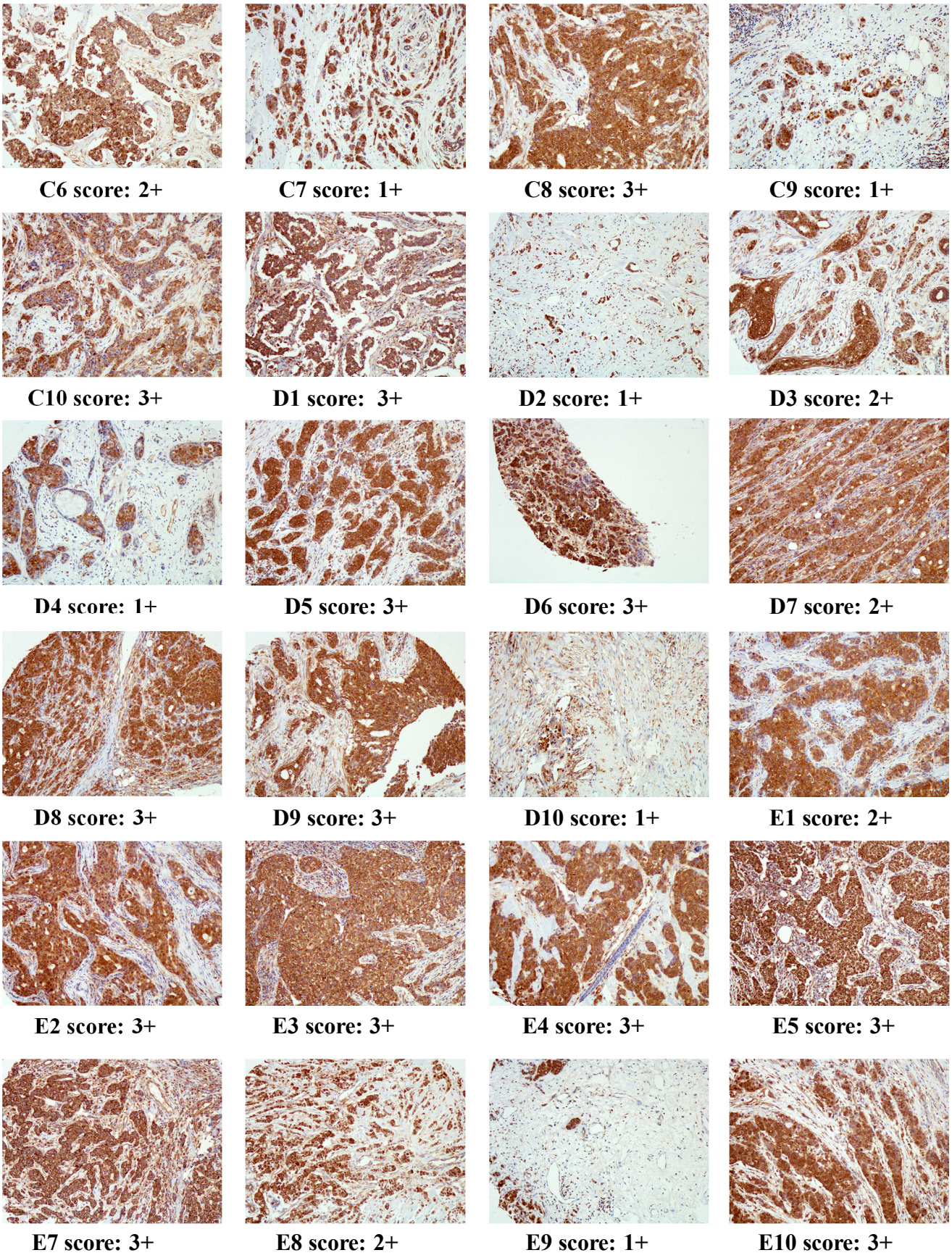
C6 score: 2+
C7 score: 1+
C8 score: 3+
C9 score: 1+
C10 score: 3+
D1 score: 3+
D2 score: 1+
D3 score: 2+
D4 score: 1+
D5 score: 3+
D6 score: 3+
D7 score: 2+
D8 score: 3+
D9 score: 3+
D10 score: 1+
E1 score: 2+
E2 score: 3+
E3 score: 3+
E4 score: 3+
E5 score: 3+
E7 score: 3+
E8 score: 2+
E9 score: 1+
E10 score: 3+

Figure S3C

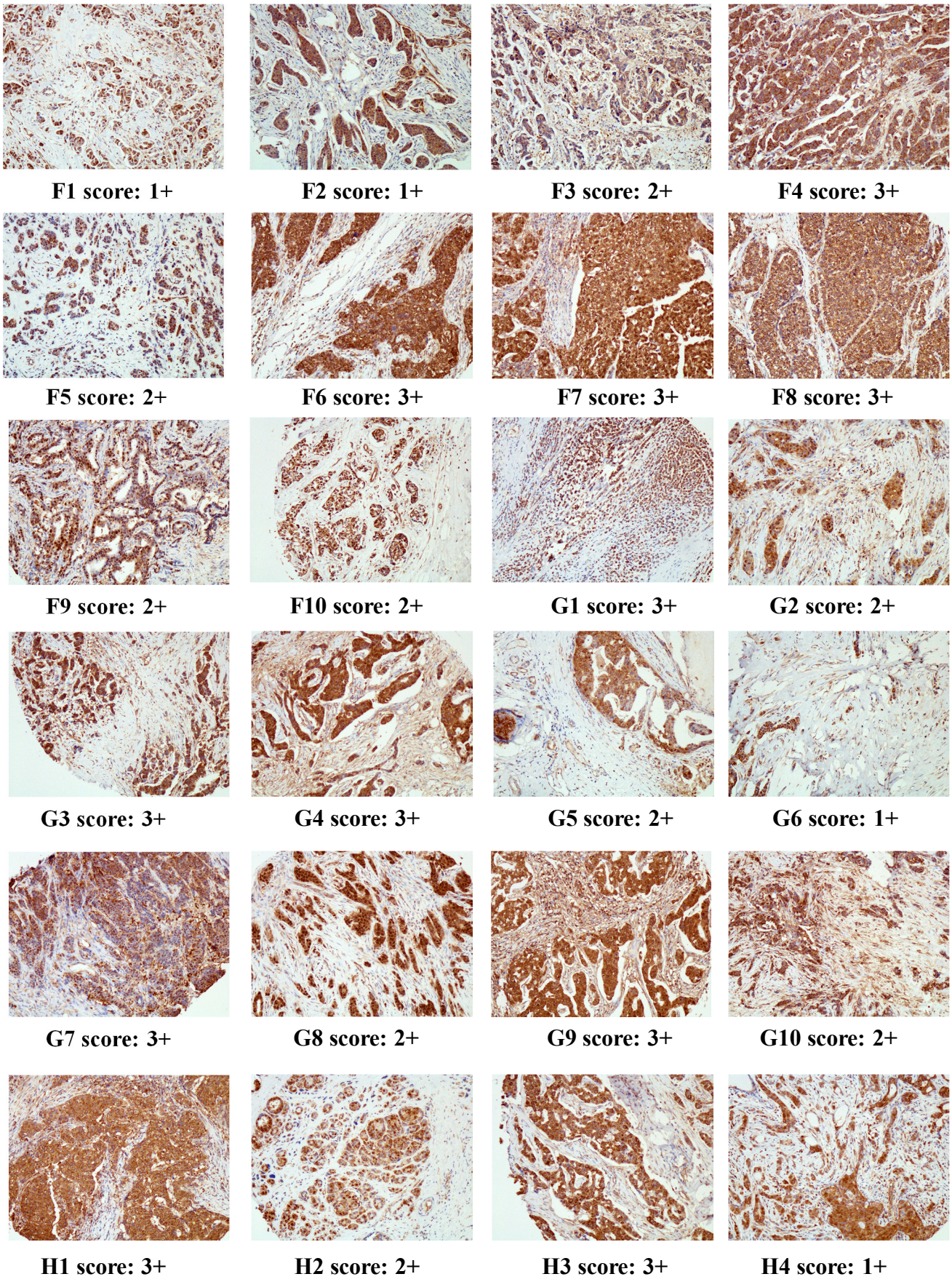
F1 score: 1+
F2 score: 1+
F3 score: 2+
F4 score: 3+
F5 score: 2+
F6 score: 3+
F7 score: 3+
F8 score: 3+
F9 score: 2+
F10 score: 2+
G1 score: 3+
G2 score: 2+
G3 score: 3+
G4 score: 3+
G5 score: 2+
G6 score: 1+
G7 score: 3+
G8 score: 2+
G9 score: 3+
G10 score: 2+
H1 score: 3+
H2 score: 2+
H3 score: 3+
H4 score: 1+

Figure S3D

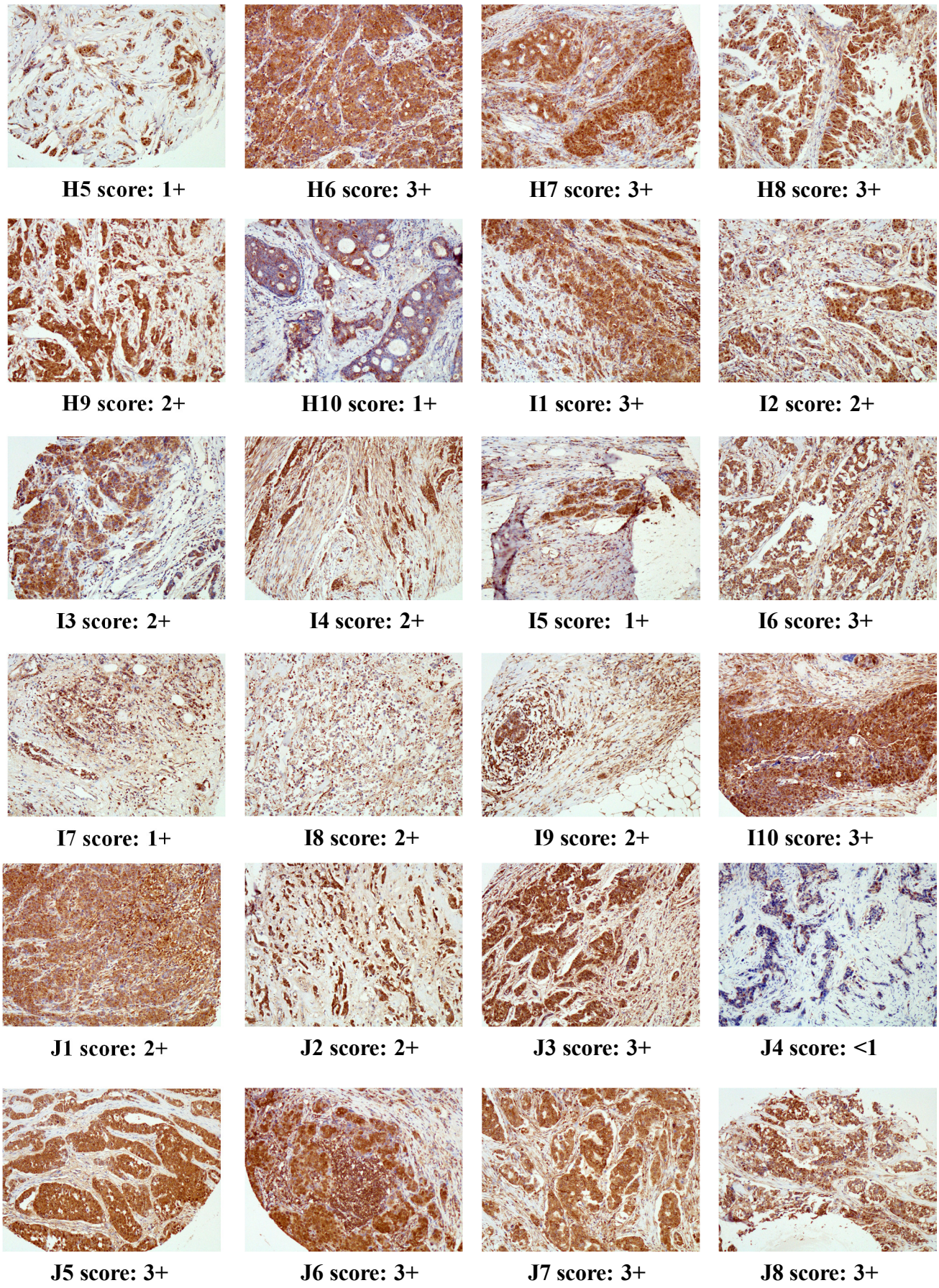
H5 score: 1+
H6 score: 3+
H7 score: 3+
H8 score: 3+
H9 score: 2+
H10 score: 1+
I1 score: 3+
I2 score: 2+
I3 score: 2+
I4 score: 2+
I5 score: 1+
I6 score: 3+
I7 score: 1+
I8 score: 2+
I9 score: 2+
I10 score: 3+
J1 score: 2+
J2 score: 2+
J3 score: 3+
J4 score: <1
J5 score: 3+
J6 score: 3+
J7 score: 3+
J8 score: 3+

Figure S3E

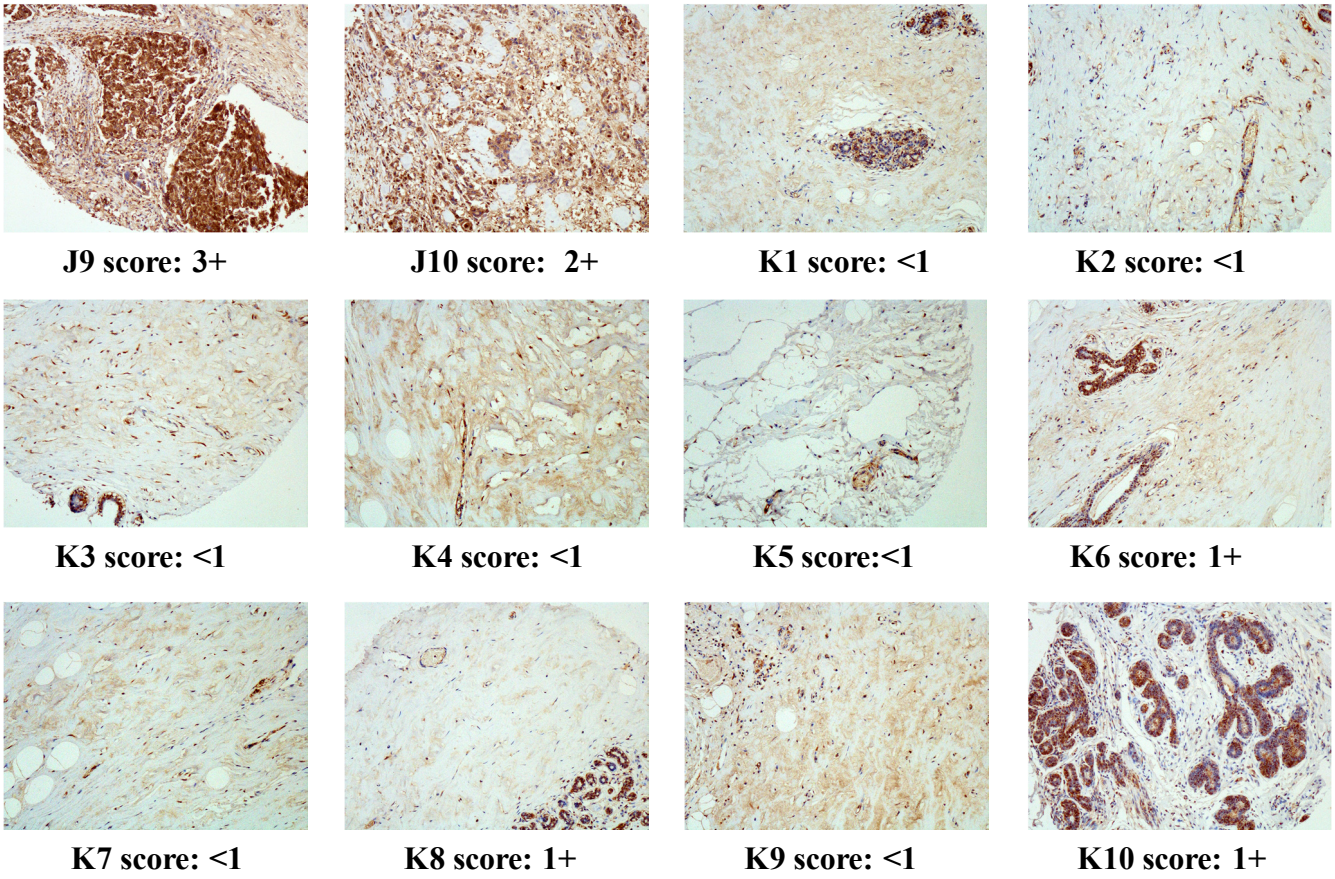

Figure S3. Immunohistochemistry staining tissue array of CPN in normal breast tissue and tumor tissue in human.

Tissue array slide contains 110 human breast tissue sections, including 100 breast cancer cases with pathological stage I-III (A1-J10) and 10 adenosis/cancer adjacent normal breast tissues as controls (K1-K10). Section B6 and E6 were missed when purchase. Detail information about the ages, pathology diagnosis, grades, stages and Tumor-Nodes-Metastasis (TNM) grading at http://www.biomax.us/tissue-arrays/Breast/BC081116a). This tissue array slide was processed by the same experimental conditions as what we used for the mouse tissue slide (see Materials and Methods). Each section was scored by a scoring system reported by Zhang et al. *(31)*.

Figure S4

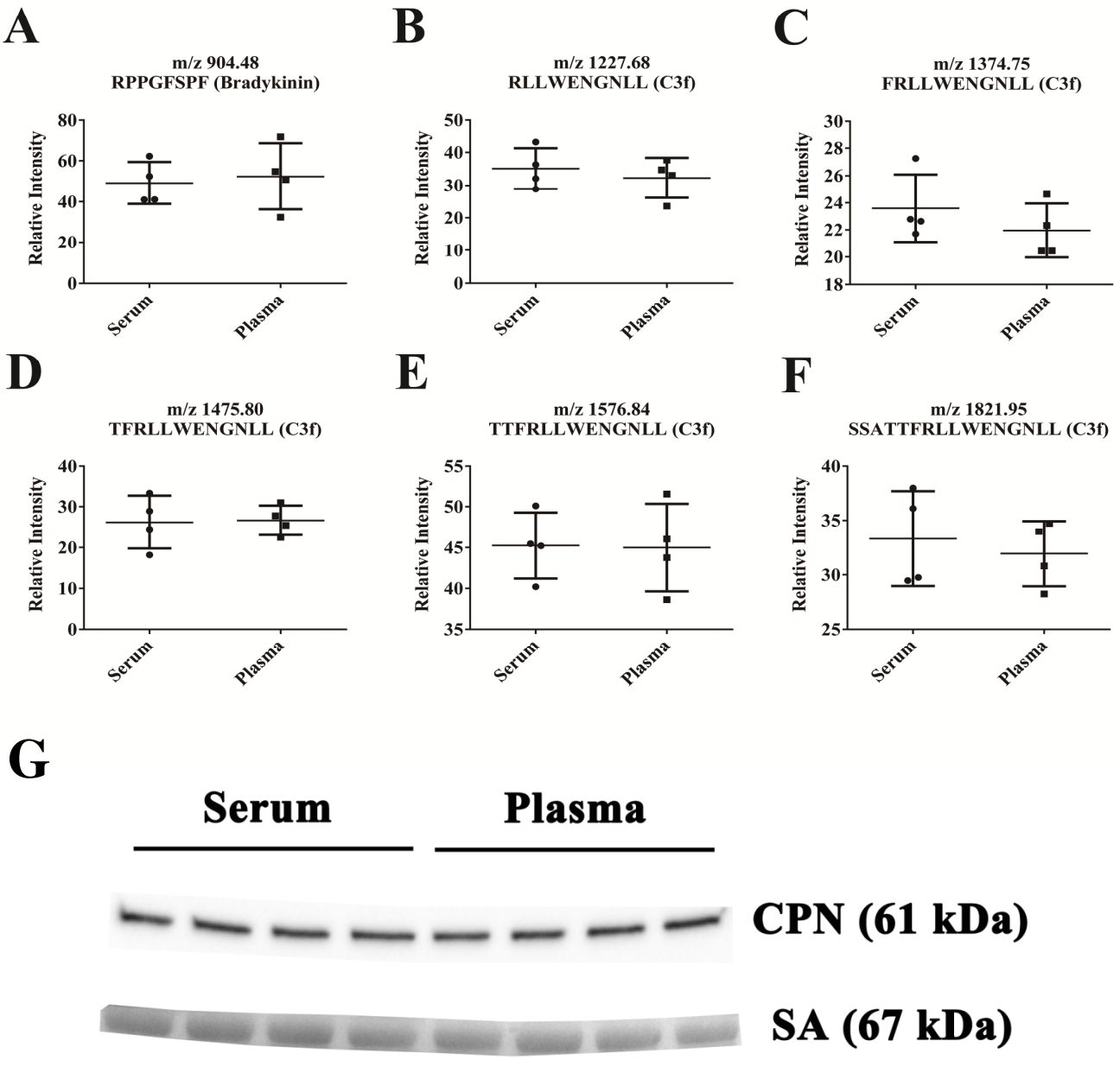

Figure S4. Comparison of the expression levels of the CPN and CPN-specific peptides in mouse serum and plasma.

Four normal nude mice (female, 6-8 weeks old) were used. **A-F.** Six CPN-specific peptides were extracted from mouse serum and plasma by nanopore fractionation, and then analyzed by MALDI-TOF MS. **G.** CPN expression by immunoblotting in mouse serum and plasma. Ponceau S-stained serum albumin serves as internal control.