Supplemental Methods

1. Clinical features of recruited infants

Signs and symptoms suggestive of clinical sepsis or NEC, and the sepsis screening procedure have been described in detail in our previous reports (Ref 12,37). In brief, the manifestations of suspected clinical sepsis/NEC included non-specific signs: *(i)* unstable temperature (< 36.5^{o}C or > 37.5^{o}C on two occasions within 12 h); *(ii)* hemodynamic instability *e.g.*, sudden increase or decrease in heart rate or persistent tachycardia (> 160/min) or bradycardia (< 100/min), poor peripheral circulation with prolonged capillary refilling time > 3 sec, systemic hypotension or unexplained increase in requirement of vasopressor support to maintain an acceptable mean arterial blood pressure; *(iii)* respiratory compromise as evidenced by progressive increase in oxygen requirement or ventilatory setting in a previously stable infant, apneic spells, sudden increase in respiratory rate or persistent tachypnea (> 60 breaths/min), and central cyanosis due to splinting of diaphragm by the distended abdomen; and *(iv)* unexplained metabolic parameters such as persistent metabolic acidosis (base deficit ≥ 10) or hyperglycemia (> 10 mmol/L). In addition, the more specific abdominal signs and symptoms, included *(i)* abdominal distension with or without tenderness, guarding or rebound; *(ii)* abdominal wall erythema; *(iii)* bile-stained gastric aspirates; *(iv)* repeated vomiting or significant regurgitation more than 2 times in 24 h; *(v)* volume of gastric residuals exceeding half of the oral intake in the previous 4 h on 2 occasions within the same day; *(vi)* bloody or blood-stained stool, and *(vii)* unexplained diarrhea.

2. Sepsis screening

All infants were recruited at the onset of clinical presentation during the initial sepsis screening, and the blood specimen (day 0) was obtained before commencement of antimicrobial treatment (Ref 12,37). During each suspected episode, a full sepsis screen was performed, which included blood, cerebrospinal fluid (CSF), and other body fluids such as peritoneal fluid, pleural fluid and pus from abscess, for microscopy and bacterial and fungal cultures. Abdominal and chest radiographs were performed, including a lateral or decubitus film if there was suspicion of bowel obstruction or

perforation. Hematologic and biochemical investigations, included *(i)* total white cell and platelet counts, *(ii)* arterial or capillary blood pH and metabolic acid-base status, *(iii)* serum and CSF glucose concentrations, and *(iv)* serial C-reactive protein (CRP) measurements. Parenteral antibiotics were immediately commenced after the infection screen had been performed.

3. Classification of clinical cases

Two distinct categories of infection episodes were prospectively defined. Group 1, the proven sepsis group, comprised infants with sepsis episodes that had been confirmed as bacterial or fungal culture-positive infection (*i.e.,* septicemia with or without meningitis or peritonitis) or NEC cases (*i.e.,* stage II or above) (Ref 40). A list of commonly identified bacterial infection in our unit is described in Supplementary Table 1 (Ref 12). Cases with positive microbial cultures from superficial body sites were likely to be colonization and were not classified as genuine infections. Group 2, the non-sepsis group, consisted of infants who met the initial screening criteria for suspected clinical sepsis but were subsequently classified as being noninfected with: *(i)* negative bacterial or fungal culture in blood, cerebrospinal fluid, urine and other sterile body fluids such as pleural effusion or ascitic fluid; *(ii)* no radiologic evidence suggestive of pneumonia or NEC; *(iii)* antimicrobial treatment terminated early, 24 – 72 h after commencement and the patient continued to improve after stoppage of medications; and *(iv)* an alternative noninfected etiology for clinical deterioration subsequently identified. In addition to Group 1 and 2, a probable clinical sepsis group (Group 3) in which suspected infants presented with at least 3 clinical signs and symptoms plus evidence of hematologic or metabolic derangements despite negative bacterial culture, was also considered to have sepsis based on strong circumstantial background of the clinical course and laboratory results.

4. Semi-quantitative proteomic profiling

The plasma proteome of each sample was enriched with hydrophobic C18 magnetic beads (Dynal, Invitrogen Corp., Carlsbad, CA) by using a KingFisher™ magnetic particle automatic processor (Thermo Fisher Scientific, Waltham, MA) and carried out in a 96-well microtiter plate platform (Ref 41). All plasma samples were assayed in a random order. The laboratory investigators were blinded

to the patients' diagnoses. Two microliters of the plasma sample were diluted and denatured with 200 μL of binding buffer (phosphate buffered saline containing 0.1% trifluoroacetic acid, v/v) and 80 μL of the diluted sample were used for each assay. During each run, 10 μL of C18 magnetic beads underwent 8 programmed steps: *(i)* activation by 100% acetonitrilte, *(ii)* washing with phosphate buffered saline, *(iii)* equilibration with binding buffer, *(iv)* sample binding for 5 min, *(v-vii)* washing with 0.1% trifluoroacetic acid for 3 times and *(viii)* elution with 40 μL of 50% acetonitrile and 0.1% trifluoroacetic acid (v/v) for 5 min. One microliter of the eluate was spotted in duplicate onto a gold ProteinChip array (Ciphergen Biosystems Inc., Fremont, CA), which was then covered with sinapinic acid matrix in 50% acetonitrile and 0.5% trifluoroacetic acid. The remaining eluate was stored at -30^{o}C for later protein identification experiments. The ProteinChip array was read on a ProteinChip PBS II reader of the ProteinChip Biomarker System (Ciphergen Biosystems) to obtain semi-quantitative proteomic profiles with a molecular mass ranging from 2 kDa to 250 kDa (Ref 41).

The spectra were smoothed and baseline subtracted with a window setting of 25 points, and externally calibrated, as previously described (Ref 16). The common peaks among the mass spectra were identified and quantified using the Biomarker Wizard software (Ciphergen Biosystems). Within the m/z range of 2000 to 20000, protein/peptide peaks were identified with both minimum valley depth and the peak height intensity greater than signal-to-noise (S/N) ratio of 3. For peak matching, a window setting of 0.3% of m/z value was used. Within the m/z range of 20000 to 250000, protein peaks were identified with both minimum valley depth and the peak height intensity greater than S/N ratio of 4. For peak matching, a window setting of 0.4% of m/z value was used. A common peak was defined when it was present in more than 10% of the cases. For those mass spectra not containing the common peak according to the peak detection definition (i.e., minimum valley depth and the peak height intensity > 3 S/N ratio for the m/z range of 2000 to 20000, > 4 S/N ratio for the m/z range of 20000 to 250000), an estimated peak at the same m/z ratio was added to complete the peak matching procedure.

Peak intensities were normalized with the total ion current, and subsequently with total peak intensities. Before data mining, normalized peak intensities of duplicate measurements were

averaged, followed by log2 transformation. Both the intra-assay and inter-assay coefficient of variations of normalized peak intensities in the majority of peaks in a control serum sample were within the range between 4% and 30% (Ref 41).

5. Denaturing non-reducing 2-D gel electrophoresis

Proteins eluted from the C18 magnetic beads was dried at 45°C by using speedvac concentrator (Eppendorf, Hamburg, Germany) and reconstituted with 185 μL of rehydration buffer (8M Urea, 2% CHAPS, 0.2% Biolyte 3-10 ampholyte, 0.001% bromophenol blue, 1mM EDTA) in the absence of reducing chemicals. An immobilized pH gradient (IPG) strip (11 cm 3-10NL, Bio-Rad Laboratories, Hercules, CA) was rehydrated with the sample overnight. For the first dimension IEF separation, the running condition was as follows: 100V for 10 min, 250V for 65 min, 500V for 25 min, 1000V for 40 min, and finally 8000V for 140 min. Second dimension SDS-PAGE was performed on 4-12% Bis-Tris polyacrylamide gels (Bio-Rad Laboratories) and the proteins were separated at 200V for 40 min in an ice bath. The 2D gel was then stained with silver nitrate by using Amersham PlusOne silver staining kit (GE Healthcare, Hammersmith Imanet, UK) with some modifications to reduce the loss of proteins with MW < 10 kDa. The gel was fixed in 40% methanol/10% acetic acid for 30 min and then sensitized with thiosulfate solution. After washing with 30% ethanol for 15 min, the gel was immersed with silver solution in 30% ethanol for 1 h. The gel was washed with Milli-Q water for 1 min for 3 times and then developed in sodium carbonate solution containing 0.03% (w/v) formaldehyde. The development was stopped by adding EDTA solution, and the gel was rinsed with Milli-Q water 3 times. Finally, gel images were obtained with GS-800 calibrated densitometer (Bio-Rad Laboratories).

6. Protein identification by mass spectrometry

The identities of the protein spots in a 2-D gel can be obtained by undertaking standard methods employing MALDI-TOF/TOF MS. The protein peaks of interest and gel spots were matched by comparing their intensities and masses. As the normalized peak intensity was directly proportional to the gel spot intensity, by comparing these two parameters in conjunction with the m/z value in mass

spectrum and the apparent molecular weight on gel, the protein spots could be matched with the protein peaks. These gel spots were then subjected to mass spectrometry analysis for their identities. The matched protein spots were excised from the silver stained gels. The gel pieces were destained, reduced with 1.75% DTT, alkylated with 350 mM iodoacetamide (IAA), and digested with modified porcine trypsin overnight (sequencing grade, Promega, Madison, WI). The tryptic peptides were harvested, cleaned up with C18 ZipTips (Millipore Corp., Billerica, MA), and subjected to MALDI-TOF/TOF MS (Ultraflex-III, Bruker Daltonics, Bremen, Germany) with α-cyano 4-hydroxy cinnamic acid as the matrix. The MS and MS/MS spectra were automatically processed with the FlexAnalysis program (version 3.0, Bruker Daltonics) with the default parameters. The MS spectrum data were searched via the online ProFound search engine to obtain the protein identity by undertaking the peptide mass fingerprinting (PMF) approach. Tandem MS data were subjected to MS/MS ion search via the Mascot search engine to obtain the protein sequence of a particular peptide. For the search parameters, 1 missed cleavage in trypsin digestion was allowed; partial oxidation of methionine, phosphorylation of serine/threonine/tyrosine, and iodoacetamide modification of cysteine residues were selected. The error tolerance values of the parent peptides and the MS/MS ion masses were 50 ppm and 0.1 Da, respectively. For a gel spot, an identification result was considered valid when both PMF and MS/MS ion search identified the same protein as the statistically significant hit from the NCBInr database, and/or when MS/MS ion search identified at least 2 tryptic peptides with sequences from the same protein as the statistically significant hits. For direct on-chip analysis, an identification result was considered valid when the MS/MS ion search returned only 1 hit with an expected-value < 0.05.

7. Protein identification by immunodepletion assay

For the proteomic peaks of m/z 8917, m/z 11731 and m/z 11918, we failed in finding the corresponding gel spots on the 2-D gels. We attempted to identify their possible identities through literature search, and then verified by immunodepletion assay. By MS analysis, Bondarenko et al (1999) showed that the observed molecular mass of human serum pro-apolipoprotein CII (pro-ApoC2) was 8914.8 Da (Ref S1), whereas Ward et al (2006) identified the SELDI peaks of m/z

11715 and m/z 11925 observed in human serum as β_2-microglobulin (B2MG) and variant of B2MG, respectively (Ref 43). For pro-ApoC2, the theoretical m/z value is m/z 8915.9, which is 0.012% different from m/z 8917. For B2MG, the theoretical m/z value of its normal form is m/z 11732.2, which is 0.010% different from m/z 11731.

To verify that the proteomic peaks of m/z 8917, m/z 11731 and m/z 11918 were pro-ApoC2, B2MG and variant of B2MG, respectively, immunodepletion assays were performed. The basic principle of an immunodepletion assay is that candidate proteins in the serum samples are first removed by using antibodies specific to their possible identities, and the depleted serum samples are then subjected to proteomic profiling. If the candidate proteins disappear specifically, the candidate proteins should have the identities as the antigens that are targeted by the antibodies. For the depletion assay of pro-ApoC2, anti-human ApoC2 antibody in 100 μL goat antiserum (Kamiya Biomedical Co., Seattle, USA) was first captured onto protein-G coated sepharose gel (75 μL gel volume, ImmunoPure Plus Immobilized Protein G, Pierce, Rockford, IL). For the control setup, protein-G coated sepharose gel was replaced with same amount of plain sepharose gel (CL-4B, GE Healthcare, Piscataway, NJ). Then 1.25 μL serum was 40-fold diluted by adding 48.75 μL PBS, added to the washed anti-ApoC2 antibody coated gel, and incubated at 4 ^{o}C overnight. After the incubation, the gel was spun down. 40 μL of supernatant, equivalent to 1 μL of ApoC2 depleted serum, was recovered, and subjected to the proteomic profiling assay. For a successful immunoassay assay, the proteomic peak of m/z 8917 should disappear in the experimental set, while it should appear in the control setup. For the depletion assay of B2MG, 1.25 μL serum was mixed with 10 uL of anti-B2MG immunoparticles, which were uniform polystyrene particles coupled with rabbit polyclonal antibody against human B2MG (DakoCytomation, Glostrup, Denmark), and topped up to a reaction volume of 50 μL with PBS. For the control setup, the anti-B2MG immunoparticle was replaced with 10 μL of plain sepharose gel (50% w/v). After overnight incubation at 4 ^{o}C, the gel was spun down. 40 μL of supernatant, which were equivalent to 1 μL of B2MG depleted serum, were recovered, and subjected to the proteomic profiling assay. For a successful immunoassay assay,

the proteomic peaks of m/z 11731 and m/z 11918 should disappear in the experimental set, while it should appear in the control setup.

8. Quantifying selected markers by immunoassay

Commercial immunoassay kits for transthyretin (DakoCytomation), β_2-microglobulin (B2MG; DakoCytomation), serum amyloid A (SAA; BioSource, Invitrogen Corp.), apolipoprotein CII (Apo-C2; LINCOplex kit, Millipore Corp.), apolipoprotein AI (Apo-A1; LINCOplex), apolipoprotein AII (Apo-A2; LINCOplex), apolipoprotein B (Apo-B; (LINCOplex), apolipoprotein CIII (Apo-C3; LINCOplex) and apolipoprotein E (Apo-E; LINCOplex) were performed according to the manufacturers' instructions.

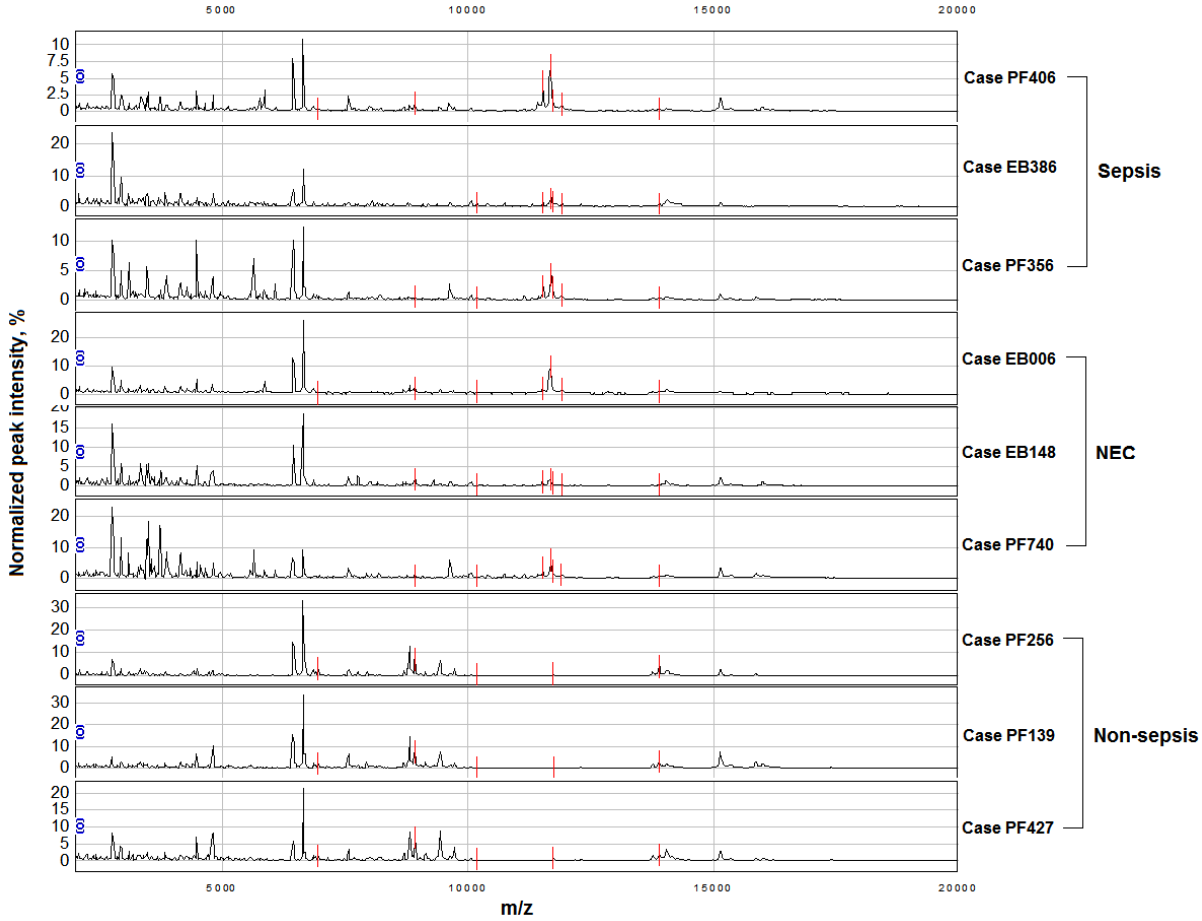
9. Additional Reference for Supplemental Method

S1. Bondarenko P.V., Cockrill S.L.,Watkins L.K., Cruzado I.D., Macfarlane R.D. 1999. Mass spectral study of polymorphism of the apolipoproteins of very low density lipoprotein. *J Lipid Res*. **40:** 543-555.

Supplemental Figures 1 – 4

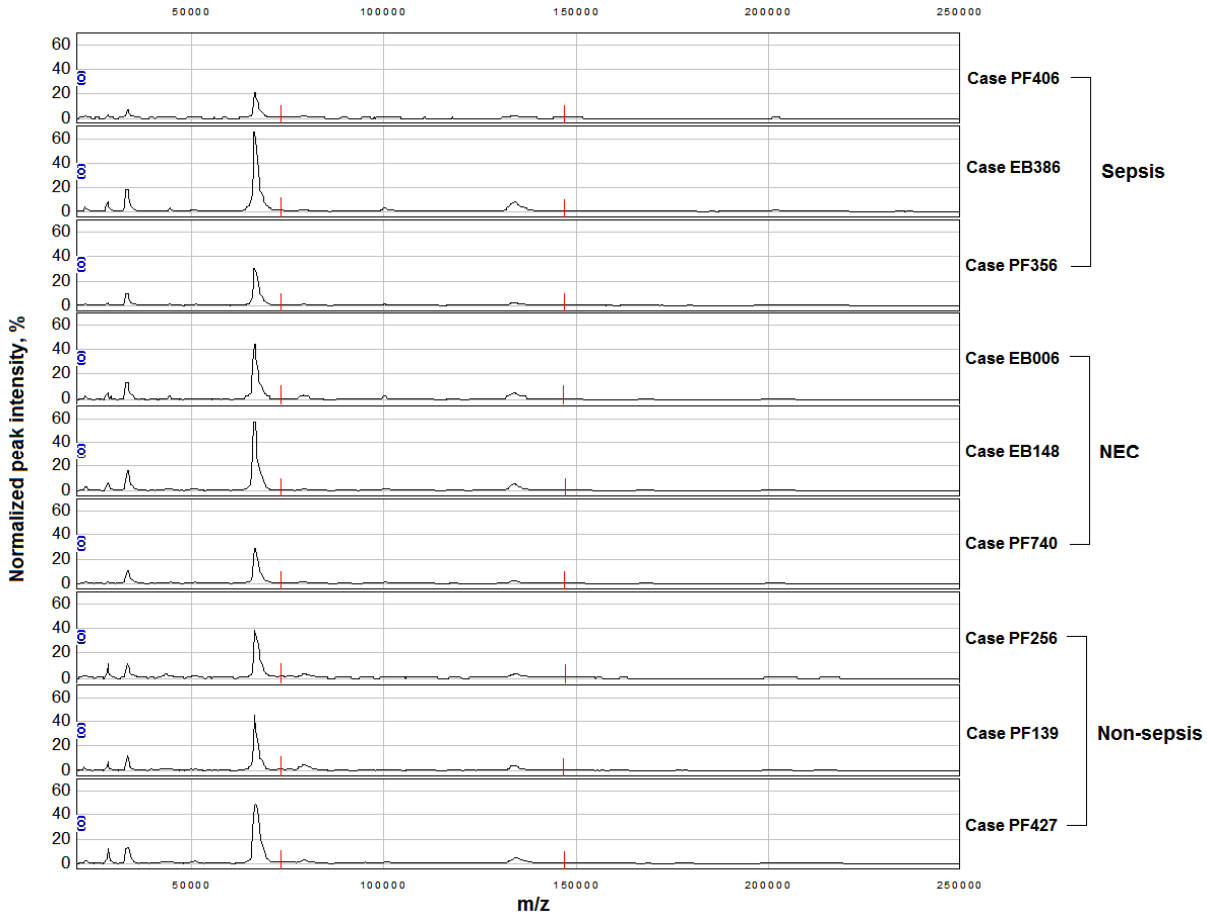
Supplemental Figure 1 (A & B): Representative plasma proteomic profiles of sepsis patients, NEC patients and non-sepsis patients. For the same set of representative cases, the zoom-in regions containing the diagnostic peaks were provided as Supplementary Figure 2 and Supplementary Figure 3.

1A. Representative plasma proteomic profiles of patients with sepsis, patients with NEC and non-sepsis patients in the m/z range of 2,000 to 20,000. The diagnostic proteomic peaks were marked with red color.

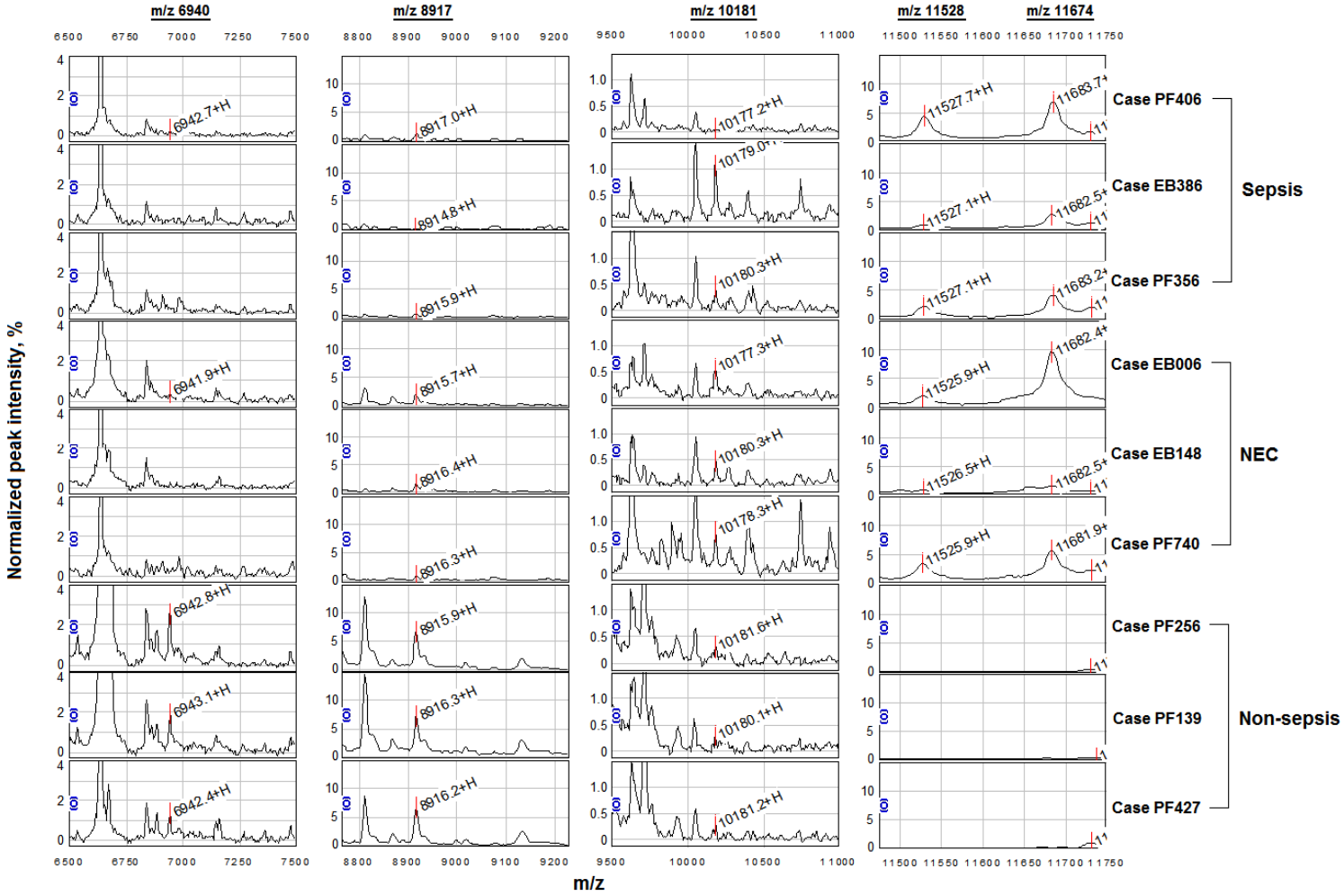


Supplemental Figure 1 (con't)

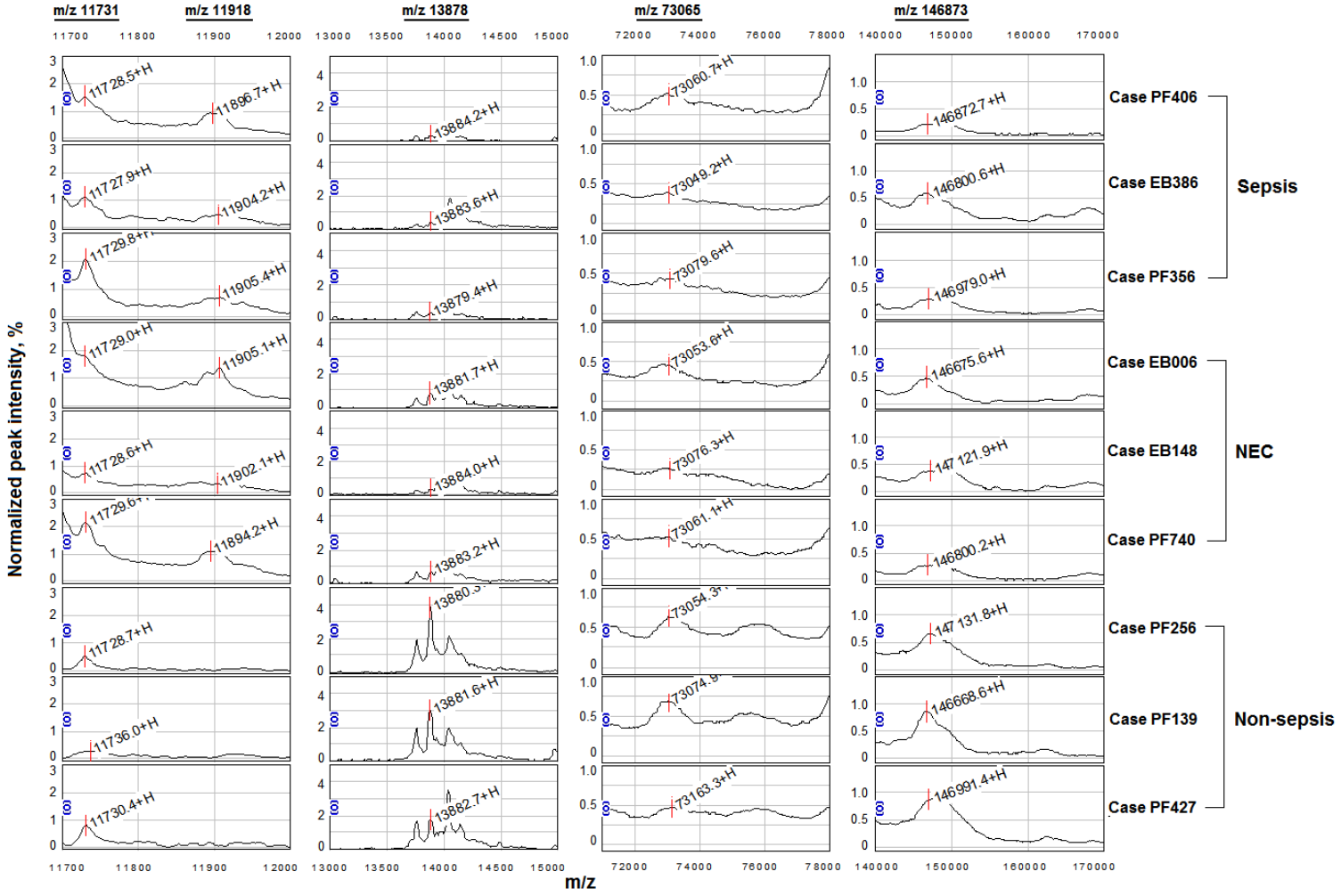
1B. Representative plasma proteomic profiles of patients with sepsis, patients with NEC and non-sepsis patients in the m/z range of 20,000 to 250,000. The diagnostic proteomic peaks were marked with red color.



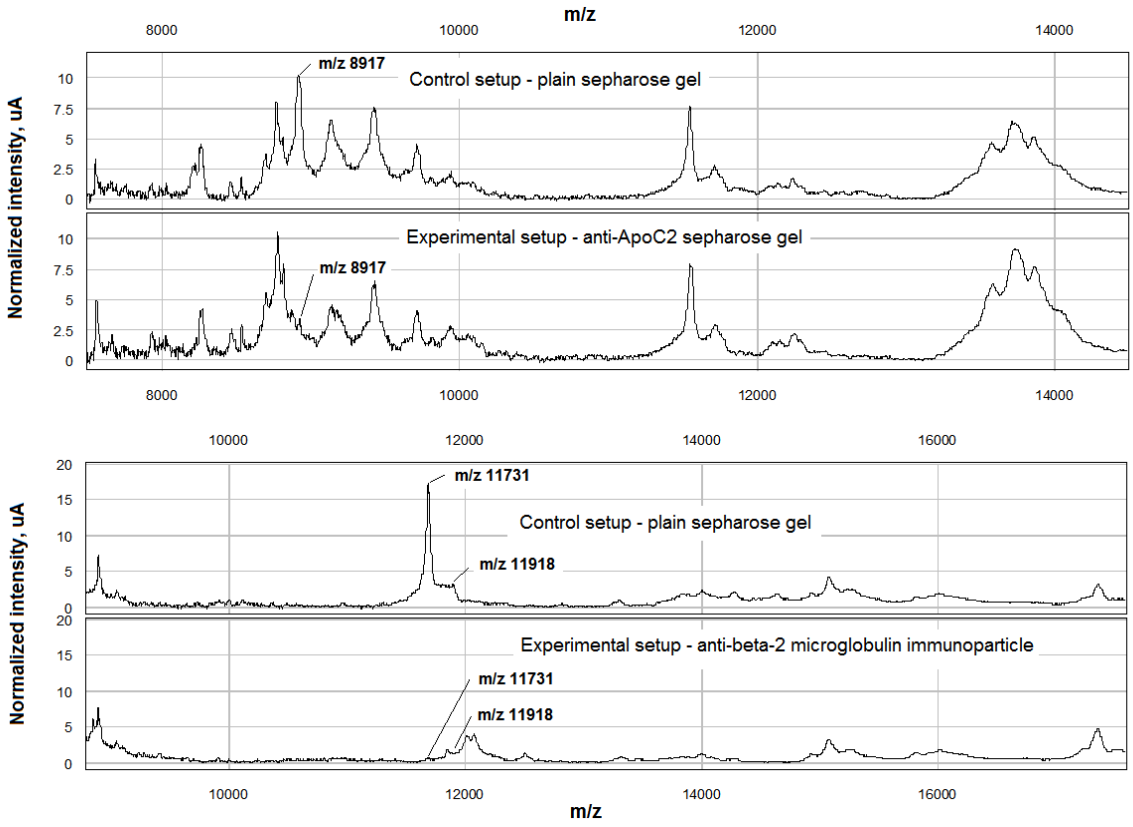
Supplemental Figure 2. Representative zoom-in regions of mass spectra of the 5 diagnostic proteomic peaks (m/z 6940, m/z 8917, m/z 10181, m/z 11528, m/z 11674) within the range of m/z 6,500 to m/z 11,700.



Supplemental Figure 3. Representative zoom-in regions of mass spectra of the 5 diagnostic proteomic peaks (m/z 11731, m/z 11918, m/z 13878, m/z 73065, m/z 146873) within the range of m/z 11,700 to m/z 250,000.



Supplemental Figure 4. Results of immunodepletion assays confirming that proteomic peaks of m/z 8917, m/z 11731 and m/z 11918 were pro-apolipoprotein CII, beta-2 microglobulin and variant of beta-2 microglobulin, respectively. By using anti-ApoC2 antibody, the peak of m/z 8917 was specifically removed, while the other peaks were unaffected. By using anti-ApoC2 antibody, the peak of m/z 8917 was specifically removed, while the other peaks were unaffected. By using anti-beta-2 microglobulin antibody, the peaks of m/z 11731 and 11918 were specifically removed, while the other peaks were unaffected.



Supplemental Tables 1 – 4

Supplemental Table 1. A list of the most common causative organisms for each category of pathogens in the neonatal unit (Ref 12).

A. Gram-positive organisms
Coagulase-negative staphylococci :
Enterococcus sp.
Staphylococcus aureus
Streptococcus bovis
Bacillus sp.
B. Gram-negative organisms :
Escherichia coli
Enterobacter sp.
Klebsiella sp.
Serratia sp.
Acinetobacter sp.
C. Fungi :
Candida albicans
Candida parapsilosis
Malassazia furfur

Supplemental Table 2. Summary of distribution patterns of the incidence of the 10 diagnostic proteomic peaks (m/z 6940, m/z 8917, m/z 10181, m/z 11528, m/z 11674, m/z 11731, m/z 11918, m/z 13878, m/z 73065, m/z 146873) detected in the biomarker discovery set comprising 37 sepsis/NEC infants and 37 non-sepsis infants. There were significantly more up-regulated diagnostic proteomic peaks detected in the sepsis/NEC cases, whereas there were significantly more down-regulated diagnostic proteomic peaks detected in the non-sepsis cases.

Distribution pattern of incidence of diagnostic proteomic peaks detected[a] in study groups	Sepsis/NEC	Non-sepsis
Number of Up-regulated proteomic peaks detected per case	***Number of cases (%)***	
5	1 (3)	0 (0)
4	10 (27)	0 (0)
3	5 (14)	0 (0)
2	12 (32)	5 (14)
1	9 (32)	31 (84)
0	0 (0)	1 (3)
	Mean (±SD)	
Average number per case	*2.5 (±1.2)*	*1.1 (±0.4)*
Mann Whitney test, 2 tailed p-value	8×10^{-6}	
Number of Down-regulated proteomic peaks detected per case	***Number of cases (%)***	
5	3 (8)	10 (27)
4	10 (27)	22 (59)
3	10 (27)	5 (14)
2	7 (19)	0 (0)
1	7 (19)	0 (0)
0	0 (0)	0 (0)
	Mean (±SD)	
Average number per case	*2.9 (±1.3)*	*4.1 (±0.6)*
Mann Whitney test, 2 tailed p-value	3×10^{-6}	

[a] For proteomic peaks within the range of 2000 m/z to 20000 m/z, the proteomic peaks with S/N ratio > 3 were detected. For proteomic peaks within the range of 20000 m/z to 250000 m/z, the proteomic peaks with S/N ratio > 4 were detected.

Supplemental Table 3. Summary of the identification results of the proteomic peaks of m/z 11528, m/z 11674, m/z 13878 and m/z 73065 by peptide mass fingerprinting and MS/MS ion search.

Proteomic peaks, m/z	Protein identity		Peptide mass fingerprinting – Profound Search Engine		MS/MS ion search – MASCOT search engine					
	Name	NCBI number	Expectation [a]	Sequence coverage, %	Observed peptide mass	Calculated peptide mass	Mass difference	Score	Expectation [a]	Peptide sequence
11528	Serum amyloid A	gi|225986	0.05	47	1455.7117	1455.7106	0.0011	51	0.008	GPGGAWAAEVISDAR
					2176.9473	2176.9562	-0.0089	84	5×10^{-6}	FFGHGAEDSLADQANEWGR
11674	Serum amyloid A	gi|225986	7×10^{-19}	46	1549.7189	1549.7202	-0.0012	64	0.0004	SFFSFLGEAFDGAR
13878	Transthyretin, Chain A	gi|55669575	4×10^{-7}	62	1365.7455	1365.7517	-0.0062	81	4×10^{-6}	GSPAINVAVHVFR
					1393.6094	1393.6150	-0.0056	58	0.0014	AADDTWEPFASGK
					2359.2490	2359.2311	0.0179	102	2×10^{-8}	YTIAALLSPYSYSTTAVVTNRK
					2450.1907	2450.1979	-0.0072	64	0.00035	ALGISPFHEHAEVVFTANDSGPR
73065	Alpha 1B-glycoprotein	gi|21071030	5×10^{-8}	18	1236.6392	1236.6390	0.0002	55	0.0025	LETPDFQLFK
					1371.6925	1371.6895	0.0030	45	0.023	HQFLLTGDTQGR
					1644.8180	1644.8181	0.0001	77	2×10^{-5}	CEGPIPDVTFELLR + carbamidomethyl (C)
					1874.9890	1874.9924	-0.0034	53	0.0027	VTLTCVAPLSGVDFQLR + carbamidomethyl (C)

[a] Expectation ≤ 0.05 indicating that the protein identification result was statistically significant ($P < 0.05$) and valid.

Supplemental Table 4. Summary of plasma apolipoprotein C2 (Apo-C2) levels, serum amyloid A (SAA) levels and Apo-SAA score in the prospective validation cohort (104 prospective episodes of suspected clinical sepsis in 60 infants).

Patient Group	Apo-C2 concentration, μg/mL mean (±SD) [number of episodes]		SAA concentration, μg/mL mean (±SD) [number of episodes]		ApoSAA score mean (±SD) [number of episodes]	
	Day 0	Day 1	Day 0	Day 1	Day 0	Day 1
Sepsis[a] /NEC	29.9 (±24.1) [42]	28.8 (±21.1) [42]	197.1 (±247.2) [42]	28.8 (±21.1) [42]	0.904 (±0.243) [42]	0.898 (±0.261) [42]
Sepsis[a] only	31.8 (±27.3) [31]	29.8 (±23.3) [31]	212.9 (±266.2) [31]	29.8 (±23.3) [31]	0.921 (±0.221) [31]	0.898 (±0.264) [31]
NEC only	24.5 (±10.7) [11]	25.9 (±13.9) [11]	152.4 (±187.4) [11]	25.9 (±13.9) [11]	0.855 (±0.304) [11]	0.896 (±0.267) [11]
Clinical probable sepsis[b]	25.0 (±14.0) [13]	26.6 (±8.3) [12]	95.9 (±143.7) [13]	26.6 (±8.3) [12]	0.929 (±0.221) [13]	0.888 (±0.273) [12]
Non-sepsis[c]	49.7 (±44.8) [49]	44.0 (±38.8) [46]	32.7 (±148.5) [49]	44.0 (±38.8) [46]	0.273 (±0.336) [49]	0.253 (±0.310) [46]
Kruskal-Wallis test, *P*-value	<0.001	0.008	<0.001	0.008	<0.001	<0.001
Dunnett's T3 post-hoc test (*P*-value)						
Sepsis/NEC vs Non-sepsis	0.027	0.070	0.001	<0.001	<0.001	<0.001
Sepsis vs Non-sepsis	0.162	0.258	0.008	0.001	<0.001	<0.001
NEC vs Non-sepsis	0.005	0.080	0.316	0.209	<0.001	<0.001
Sepsis vs NEC	0.759	0.984	0.955	0.885	0.982	1.000
Probable clinical sepsis vs Non-sepsis	0.005	0.021	0.432	0.311	<0.001	<0.001
Probable clinical sepsis vs Sepsis	0.850	0.984	0.332	0.156	1.000	1.000

[a] Proven sepsis with positive blood culture.

[b] Probable clinical sepsis with which suspected infants presented with at least 3 clinical signs and symptoms plus evidence of hematologic or metabolic derangements despite negative bacterial culture.

[c] Non-sepsis cases who were suspected clinical sepsis but were subsequently classified as being noninfected and had a definitive diagnosis unrelated to sepsis or NEC.