

# Northumbria Research Link

Citation: Stewart, Christopher, Nelson, Andrew, Treumann, Achim, Skeath, Tom, Cummings, Stephen, Embleton, Nicholas and Berrington, Janet (2015) Metabolomic and proteomic analysis of serum from preterm infants with necrotising enterocolitis and late-onset sepsis. *Pediatric Research*, 79 (3). pp. 425-431. ISSN 0031-3998

Published by: Nature Publishing Group

URL: <http://dx.doi.org/10.1038/pr.2015.235> <<http://dx.doi.org/10.1038/pr.2015.235>>

This version was downloaded from Northumbria Research Link:  
<http://nrl.northumbria.ac.uk/25405/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

[www.northumbria.ac.uk/nrl](http://www.northumbria.ac.uk/nrl)



Open

# Metabolomic and proteomic analysis of serum from preterm infants with necrotising enterocolitis and late-onset sepsis

Christopher J Stewart<sup>1</sup>, Andrew Nelson<sup>1</sup>, Achim Treumann<sup>2</sup>, Tom Skeath<sup>3</sup>, Stephen P Cummings<sup>1</sup>, Nicholas D Embleton<sup>3</sup> and Janet E Berrington<sup>3</sup>

**BACKGROUND:** Necrotising enterocolitis (NEC) and late-onset sepsis (LOS) are the leading causes of death among preterm infants in the developed world. This study aimed to explore the serum proteome and metabolome longitudinally in preterm infants with NEC or LOS, matched to controls.

**METHODS:** Nineteen patients (10 cases, 9 controls) were included. A sample 14 d prior to and following, as well as at disease diagnosis, was included for cases. Controls had serum matched at diagnosis for corresponding case. All samples ( $n = 39$ ) underwent shotgun proteomic analysis, and 37 samples also underwent metabolomics analysis using ultra performance liquid chromatography–tandem mass spectrometry.

**RESULTS:** The proteomic and metabolomic profiles of serum were comparable between all infants. Eight proteins were associated with NEC and four proteins were associated with LOS. C-reactive protein was increased in all NEC patients at diagnosis.

**CONCLUSION:** No single protein or metabolite was detected in all NEC or LOS cases which was absent from controls; however, several proteins were identified which were associated with disease status. The differing expression of these proteins between diseased infants potentially relates to differing pathophysiology of disease. Thus, it is unlikely a single biomarker exists for NEC and/or LOS.

**N**ecrotising enterocolitis (NEC) and late-onset sepsis (LOS) remain important yet potentially preventable causes of death and serious disability in preterm infants (1). Recent comprehensive studies of gut bacterial communities have either failed to demonstrate or produced inconsistent results, with respect to identifying a potential causative organism or reproducible differences in gut bacterial profiles (2–4). In part, this may be due to the multifactorial etiology of NEC and also reflect the challenges in making a clinical diagnosis when surgery (and therefore histological confirmation) is not required (5). In addition, LOS frequently presents with abdominal distension and raised inflammatory markers, yet blood cultures may be negative meaning that distinguishing between NEC and LOS is challenging. Even though both diseases may involve perturbations in gut bacterial communities, preventative measures and treatments differ.

Longitudinal noninvasive (e.g., stool/urine) sampling from preterm infants offers the potential to explore disease mechanisms and identify potential biomarkers. In addition, blood can be successfully salvaged from preterm infants after routine clinical tests are complete (6). Recent developments in mass spectrometry (MS) have facilitated shotgun proteomics on complex samples which rely on MS to detect peptides (digested protein) in a given sample, overcoming limitations of gel-based approaches. Metabolomics is the detection of small molecules (metabolites), which are the functional end point of cellular processes. Unlike genomics and proteomics, the accurate identification of metabolites requires confirmation with known standards and thus remains a major limitation of this emerging technology.

Previous work in preterm biomarker research has used targeted approaches, identifying many potential biomarkers including intestinal fatty acid-binding protein, serum amyloid A, and C-reactive protein that have been used alone or in combination, but with little impact to date on clinical care (7–9). A recent metabolomic pilot study showed that changes in the serum metabolome prior to NEC are associated with upregulation of interleukin-1 $\beta$  (IL-1 $\beta$ ), although this study involved only five patients (10 samples) with NEC and sample timing in relation to NEC was quite varied (10). A two-dimensional gel proteomic study of intestinal tissue from premature piglets with NEC compared with controls identified 19 differentially expressed proteins with functions involving oxidative stress, signal transduction, protein folding and degradation, oxygen transport, and energy metabolism (11).

The aim of our study was to use proteomic and high-throughput metabolomic analysis on prospectively collected serum from preterm infants diagnosed with NEC or LOS. We analyzed longitudinal samples in diseased patients, compared with controls matched for gestation, to determine differences in the serum proteome and metabolome that might be predictive or indicative of disease.

## RESULTS

In total, 447 unique proteins and 24,153 metabolites (16,882 positively ionised and 7,271 negatively ionised) were detected. Known and potentially relevant standards were ran alongside

<sup>1</sup>Faculty of Health and Life Sciences, Northumbria University, Newcastle upon Tyne, UK; <sup>2</sup>Newcastle University Protein and Proteome Analysis, Newcastle University, Newcastle upon Tyne, UK; <sup>3</sup>Newcastle Neonatal Service, Royal Victoria Infirmary, Newcastle upon Tyne, UK. Correspondence: Christopher Stewart ([christopher.stewart@northumbria.ac.uk](mailto:christopher.stewart@northumbria.ac.uk))  
Received 2 June 2015; accepted 22 August 2015; advance online publication 9 December 2015. doi:10.1038/pr.2015.235

the metabolomic samples, including short-chain fatty acids, ceramides, and amino acids but could not be accurately matched to identified metabolites from samples. The most abundant proteins detected included  $\alpha$ -2-macroglobulin,  $\alpha$ -1-antitrypsin, serotransferrin, complement C3, and fibrinogen  $\alpha$  and  $\beta$  chain. Despite depletion of serum albumin and IgG, serum albumin was still found to be the seventh most abundant protein based on relative intensity. Proteins associated (i.e., increased in relative intensity) with early samples (<20 day of life (DOL)) included angiotensinogen,  $\alpha$ -fetoprotein, and antithrombin-III, whereas the proteins associated with the older samples (>21 DOL) included histidine-rich glycoprotein, C4b-binding protein  $\alpha$  chain, calcium-independent phospholipase, hemoglobin subunit  $\beta$ ,  $\alpha$ -1-acid glycoprotein 1, and leucine-rich  $\alpha$ -2-glycoprotein.

As expected, partial least squared discriminatory analysis of disease samples matched to controls showed low R2 and Q2 scores due to the overall comparability of serum samples from disease and control samples (Figure 1). However, the proteomic and metabolomic data were in concordance and demonstrated that serum from patients 180 and 174, both of whom were diagnosed with NEC and underwent surgery, were distinct from other serum profiles within this study. The protein most associated with these samples was identified as transforming growth factor beta induced protein. Notably, patient 180 represents the biggest outlier, particularly in the metabolomics profiling, and was the only patient who died in this cohort.

Samples from both diseased and healthy individuals were comparable, and there were no unique proteins or metabolites consistently found in only samples from patients with NEC or LOS. Supervised modeling of the data revealed that the expression of several proteins increased in patients diagnosed with NEC and LOS, compared with controls (Supplementary Figure S1 online). Eight proteins associated with NEC and four proteins associated with LOS in both the between-control analysis (phase I) and the within-patient analysis (phase II) are listed in Table 1. Notably, the relative abundance of these proteins was not increased in all patients at diagnosis, with some patients showing negative correlations and a relative abundance lower than the average of all matched controls (Figure 2). C-reactive protein was increased in all NEC patients at diagnosis and exceeded the average for controls. Ig  $\alpha$ -2 chain C region levels were increased in 5 out of 6 NEC patients at diagnosis (compared to 14 d prior to diagnosis) and had a higher relative abundance than controls in 4 out of 6 patients. Macrophage migration inhibitory factor was also increased in 3 out of 6 patients with NEC at diagnosis and was above the average control level in 4 out of 6 patients at diagnosis.

## DISCUSSION

### Serum Proteome in the Early Stages of Life

$\alpha$ -2-Macroglobulin and  $\alpha$ -1-antitrypsin were the two most abundant proteins detected in this study, which is in agreement with previous findings from preterm infants in the first 6 mo of life (12). In accordance with the findings of this study,

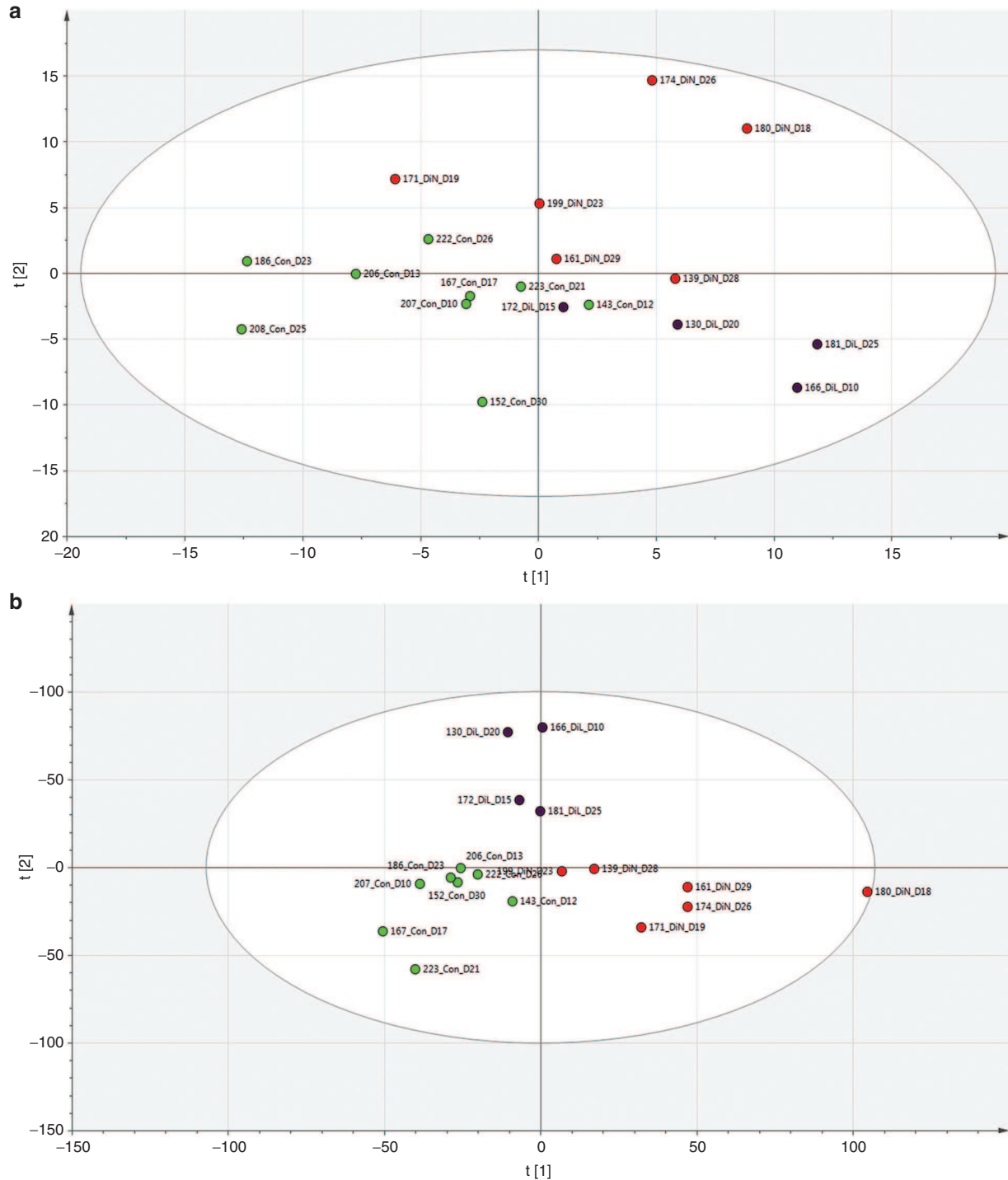
a historical targeted *in vivo* investigation showed no clear association of these two proteins in preterm neonates with poor health (13). Similarly,  $\alpha$ -fetoprotein levels reduce from week 14 of gestation although the synthesis of this protein is known to continue following birth (14). Thus, the association of  $\alpha$ -fetoprotein with early samples likely relates to the prematurity of this cohort as suggested by other studies (15).

The proteins associated with samples analyzed later in life demonstrate the increasing complexity of the serum proteome with age. C4b-binding protein  $\alpha$ -chain, which inhibits elements of the complement system, is increased in later samples potentially representing increased exposure of the host to potential pathogens (2). Like the proteins associated with the early samples, the proteins associated with the later samples are also in accordance with published data, such as  $\alpha$ -1-acid glycoprotein 1 which increases from infancy to adolescence (16). This agreement with earlier studies suggests that the shotgun metaproteomic methodology developed in this study is valid.

### Serum Proteome and Metabolome in Health and Disease

Profiling of the serum proteome and metabolome are powerful techniques for determining functional changes within a host in response to disease. The analysis was split into two phases to identify the most robust and reproducible proteins and metabolites associated with disease. Phase I explored increased expression between diseased and gestationally matched control infants, and phase II identified those increased within patients at disease diagnosis (Figure 3). The serum proteome and metabolome were found to be highly conserved within and between patients. Thus, no distinct profile was found which could robustly predict patients diagnosed with NEC or LOS. Furthermore, no unique proteins or metabolites were found only in samples representing disease or health. Similarly, in a comparable cohort where the urine metabolome was investigated, no clear difference between NEC and control infants was found using ordination analysis, and subsequently, no distinct urinary metabolite biomarker that could predict the onset on NEC (17).

Eight upregulated proteins were associated with samples at NEC diagnosis and four different proteins were associated with LOS diagnosis. In accordance with previous studies, upregulation of C-reactive protein and serum amyloid A were associated with the diagnosis of NEC (7–9,18). Here, C-reactive protein was upregulated in all NEC patients at diagnosis and exceeded the average for controls. This likely demonstrates increased inflammation in infants who developed NEC (9). Using a shotgun metaproteomic approach also facilitated the identification of novel proteins not previously reported in association with NEC and LOS, including Ig  $\alpha$ -2 chain C region and macrophage migration inhibitory factor. Increased expression of these proteins may relate to increased immune response against foreign antigens contributing to NEC (19). It is important to note that in some cases the expression of proteins in patients diagnosed with disease was lower than the average for the control patients, suggesting that changes of the serum proteome in patients with NEC or LOS are specific to



**Figure 1.** Partial least squared discriminatory analysis score scatter plot of serum protein and metabolite profiles. DiN show samples at NEC diagnosis (red), DiL show samples at LOS diagnosis (purple), and Con show corresponding matched control samples (green). (a) Proteomics.  $R^2 = 0.56$ ,  $Q^2 = -0.15$ . (b) Metabolomics.  $R^2 = 0.73$ ,  $Q^2 = -0.21$ . LOS, late-onset sepsis; NEC, necrotising enterocolitis.

the individual or a result of different factors cascading to NEC pathogenesis (20). Notably, both the protein and metabolite profiles of the only infant to die in this study (patient 180) were distinct from the other disease and control infants.

Compared with existing studies in comparable cohorts, we find commonality between potentially useful biomarkers for the diagnosis of preterm disease; however, no single serum

biomarker is specific enough to robustly separate the clinical manifestation of NEC from controls. The stool microbiota has been shown to have important influences on NEC and LOS, but the stool microbiota varies vastly within and between infants and the “causative organism” differs between studies (2–4). Thus, while robust exploration of the direct influence of the gut microbiota on the serum proteome is important, statistical

**Table 1.** Proteins associated with diagnosis of disease

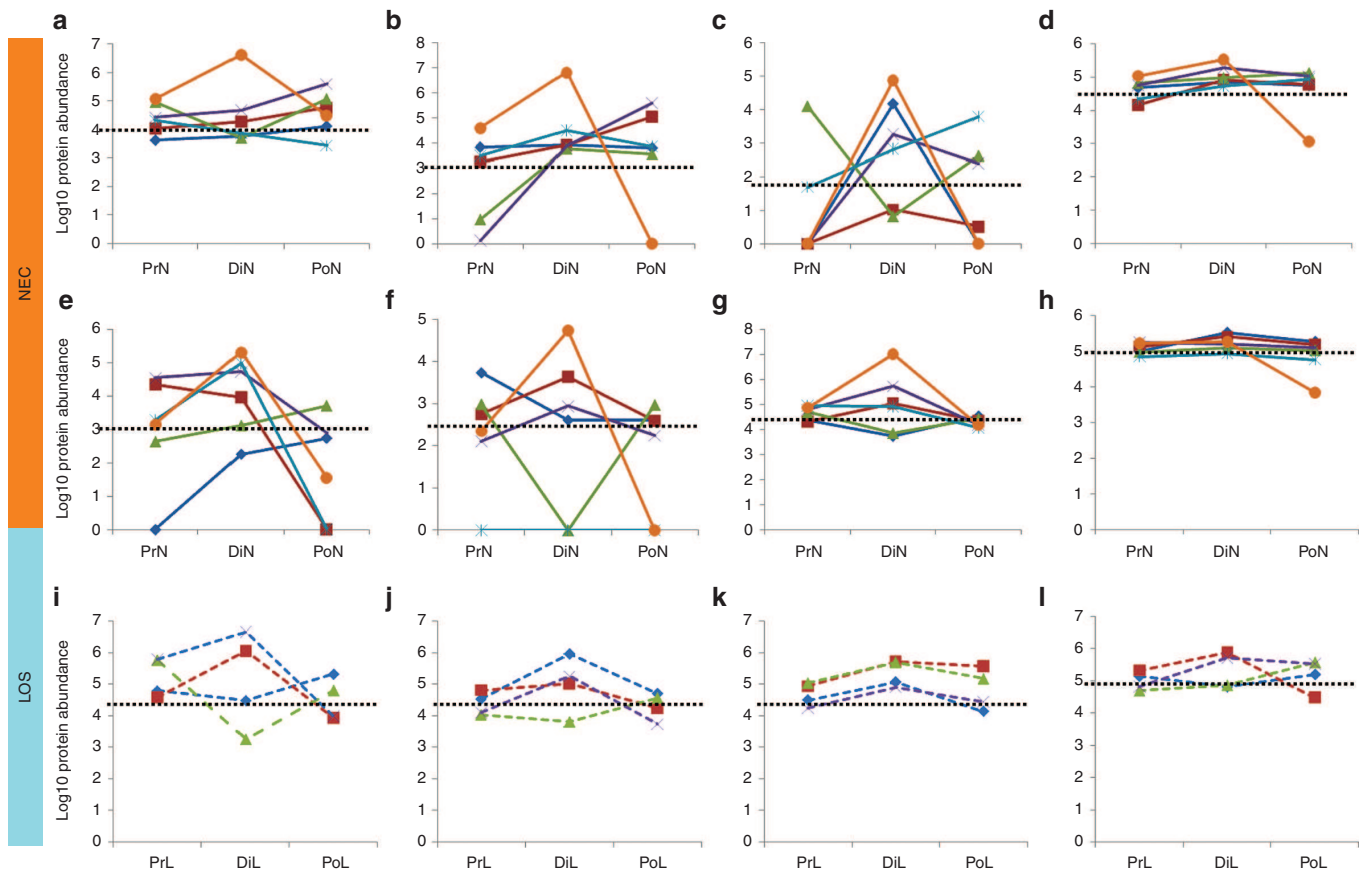
	Accession	Protein identification
NEC diagnosis associated	P02741	C-reactive protein
	Q5VVP7	C-reactive protein (1–205)
	A0A075B6N7	Ig $\alpha$ -2 chain C region (fragment)
	P07355-2	Isoform 2 of annexin A2
	P05451	Lithostathine-1- $\alpha$
	P14174	Macrophage migration inhibitory factor
	P0DJ19	Serum amyloid A-2 protein
	Q15582	Transforming growth factor- $\beta$ -induced protein ig-h3
LOS diagnosis associated	A0A087WU08	Haptoglobin
	P23634-4	Isoform XK of plasma membrane calcium-transporting ATPase 4
	A0A087WV45	Transthyretin
	O75643	U5 small nuclear ribonucleoprotein 200 kDa helicase

LOS, late-onset sepsis; NEC, necrotising enterocolitis.

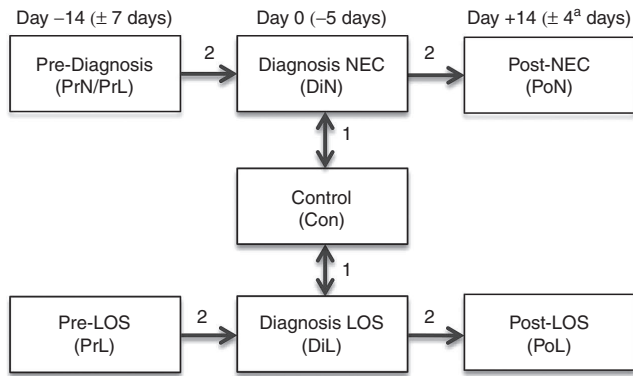
validity will require large numbers of infants and time points. It is noteworthy that local gut protein and small-molecule biomarkers in stool may also offer important insights into disease pathophysiology; however, to control for proteases and extract high-quality stool protein from a functioning neonatal intensive care unit is outside the scope of this study.

**Limitations and Future Research Direction**

We employed longitudinal sampling at discrete times relative to disease to identify new targets that might be used to facilitate future biomarker discovery or provide new insights into disease pathology. It is however noteworthy that this study alone is not sufficiently powered to determine robust biomarkers for use in the clinical diagnosis of NEC and LOS. This proof-of-concept study also utilized serum that was frozen after salvage following routine clinical work, which may result in degradation of some proteins. Nevertheless, we feel that such opportunistic approaches will be necessary if research in NEC and sepsis is to be progressed, especially given the ethical challenges posed by invasive sampling and our current inability to accurately predict individuals who will develop disease.



**Figure 2.** Line graph of longitudinal change within patients with NEC and LOS. Plots show most associated proteins in accordance with Table 2. Protein abundance transformed with Log10. NEC samples shown by solid lines in panels (a–h): patients 139 (dark blue diamond), 161 (red square), 171 (green triangle), 180 (purple cross), 199 (light blue cross), 174 (orange circle). LOS samples shown by dashed lines in panels (i–l): patients 130 (dark blue diamond), 166 (red square), 172 (green triangle), 181 (purple cross). Black dotted line represents the average of control data. Each panel represents a different protein: (a) C-reactive protein, (b) C-reactive protein (1–205), (c) Ig  $\alpha$ -2 chain C region, (d) isoform 2 of annexin A2, (e) lithostathine-1- $\alpha$ , (f) macrophage migration inhibitory factor, (g) serum amyloid A-2 protein, (h) transforming growth factor- $\beta$ -induced protein ig-h3, (i) Haptoglobin, (j) isoform XK of plasma membrane calcium-transporting ATPase 4, (k) transthyretin, (l) U5 small nuclear ribonucleoprotein 200 kDa helicase. LOS, late-onset sepsis; NEC, necrotising enterocolitis.



**Figure 3.** Study design. Numbers on arrows represent the phase of analysis, where phase I determined features associated with disease diagnosis compared to control samples and phase II determined features associated with disease diagnosis temporally within patients. <sup>a</sup>Last available sample from patient 139 was 3 d post diagnosis.

This study developed an optimized powerful methodology for the functional profiling of serum using both proteomics and metabolomics. There is great potential for metabolomics to detect small molecules in serum, potentially related to functional changes of the host in response to disease. Metabolomics still offers important advances in disease understanding for independent identification of significant metabolites, even when used solely as a profiling technology. Proteomic technologies offer the advantage of identifying proteins from existing databases based on peptide matches. Future work should build on the findings here, incorporating longitudinal sampling in greater numbers of both disease and control patients. Combining multi-omic datasets into systems biology analysis will facilitate increased understanding of changes in pathways and provides a means of cross-validation between datasets.

## Conclusion

No single protein or metabolite was detected in all NEC or LOS cases which was absent from controls; however, several proteins were identified which were associated with disease status. The expression of these proteins generally varied between diseased infants, potentially relating to differing pathophysiology of disease. Thus, it is unlikely a single biomarker exists for distinguishing NEC and/or LOS from healthy infants. Future work should aim to better understand the multiple mechanisms that lead to NEC or LOS, allowing more accurate classification of patients into disease type. This will prevent confounding analysis based on inaccurate disease classification and may allow more specific biomarkers to be developed, accurately predicting the health status of an individual.

## METHODS

### Ethics Statement

Ethical approval was obtained from the County Durham and Tees Valley Research Ethics Committee, and signed informed parental informed was obtained.

### Infants and Samples

All infants were cared for in the neonatal intensive care unit of the Royal Victoria Infirmary, Newcastle upon Tyne, UK, between November

2011 and August 2012. Standardized feeding, antibiotic, and antifungal guidelines were used as described previously (2). Samples were salvaged and stored at  $-20^{\circ}\text{C}$  within 72 h as previously described (6). “Cases” had confirmed NEC and/or LOS categorized independently by the attending clinician and blindly confirmed (J.E.B. or N.D.E.) as described previously (21). Ten cases were matched with 10 controls based on gestational age, although issues in the preparation of serum arose in one control case, leaving nine control infants (Table 2). For cases, a sample 14 d ( $\pm 7$ ) prior to disease diagnosis, immediately prior to day of diagnosis ( $\pm 5$ ), and 14 d ( $\pm 4$ ) after disease diagnosis were analyzed, with the exception of patient 139 in whom the last available sample was 3 d post diagnosis (Table 2 and Figure 1). Control serum was chosen as close to the day of life of the case disease onset as available ( $\pm 6$  d). All samples ( $n = 39$ ) underwent proteomic analysis, and 37 samples also underwent metabolomics analysis, due to insufficient sample remaining for two samples.

### Serum Proteomics

A 10  $\mu\text{l}$  aliquot of serum was depleted using the Top 2 abundant protein depletion spin columns (Pierce, Rockford, IL) for the removal of albumin and IgG according to the manufacturers’ instructions. The protein concentration was then determined in triplicate using the Micro BCA protein assay kit (Pierce) and adjusted to 5  $\mu\text{g}$  in 50  $\mu\text{l}$  with 25 mmol/l ammonium bicarbonate in siliconized microfuge tubes. A 1:1 v/v of 50% sodium deoxycholate was added and incubated on ice for 5 minutes to denature the protein (22). The denatured protein was diluted in 25 mmol/l ammonium bicarbonate and reduced with 50 mmol/l dithiothreitol at  $60^{\circ}\text{C}$  for 30 min. The sample was then cooled to room temperature and alkylated with 100 mmol/l iodoacetamide at room temperature in the dark for 30 min. To digest the protein, 5.8  $\mu\text{g}/\text{ml}$  of trypsin (Promega, Madison, WI) was added and incubated overnight ( $\sim 16$  h) at  $37^{\circ}\text{C}$  to give a 1:20 trypsin:protein ratio. Following the overnight incubation, half the amount of trypsin was added and incubated for a further 2 h to promote further digestion. The reaction was stopped by the addition of 10% v/v formic acid, and the sample was centrifuged at 16,000g to pellet sodium deoxycholate. The sample was freeze-dried and stored at  $-80^{\circ}\text{C}$  until further processing.

A multistep liquid chromatography (LC) gradient was used with 5% acetonitrile (ACN) increasing to 40% ACN over 180 min, with a further 90% ACN for 20 min, followed by a final 10 min re-equilibration at 5% ACN. Samples were run in triplicate, and the order of samples in each triplicate sequence was randomized. An Acclaim PepMap RSLC C18, 2  $\mu\text{m}$ , 100  $\text{\AA}$  50 cm Easy-Spray column (Thermo-Scientific, Waltham, MA) was used for peptide separation. A Q-Exactive (Thermo-Scientific) was used for top 10 MS/MS.

### Serum Metabolomic Profiling

Water, methanol, and ACN were liquid chromatography mass spectrometry (LCMS) grade (Sigma-Aldrich, St Louis, MO). Metabolites were extracted from 30  $\mu\text{l}$  serum and homogenized in 120  $\mu\text{l}$  cold 100% methanol by vortexing for 15 min at  $4^{\circ}\text{C}$ . The suspension was then centrifuged at 10,000g for 10 min at  $4^{\circ}\text{C}$  and lyophilized in a freeze dryer before storage at  $-80^{\circ}\text{C}$ . Samples were re-suspended in 120  $\mu\text{l}$  of initial start phase buffer (5% ACN). Where 30  $\mu\text{l}$  of sample was unavailable, the total volume was recorded, and the volume of initial start phase buffer was adjusted accordingly. Serum metabolite profiling was performed using reverse-phase ultraperformance LCMS tandem mass spectrometry. A 2.6  $\mu\text{m}$ ,  $150 \times 2.1$  mm Accucore C18 column (Thermo-Scientific) was used at  $40^{\circ}\text{C}$  with a 3.0  $\mu\text{l}$  injection and 300  $\mu\text{l}/\text{min}$  flow rate throughout. A multistep LC gradient was used with 5% ACN increasing to 95% ACN over 22 min, with a further 95% ACN for 4 min followed by a final 4 min re-equilibration at 5% ACN. Samples were run in triplicate, and the order of samples in each triplicate sequence was randomized. A blank consisting of LCMS-grade water underwent the same procedure, and an aliquot of every sample was used as a pool. Prior to each run, a blank and five pools were processed to equilibrate the system, then blanks and pools were processed periodically every 10 samples for background subtraction and quality control, respectively. A Q-Exactive (Thermo-Scientific) was used for the MS, and metabolomic profiling was performed using HESI with high-resolution (70,000) positive and negative switching. The mass range was set from 100 to 1,000 m/z.

**Table 2.** Patient demographics

	Patient number	Gestational age (wk)	Birth weight (g)	Delivery mode	Feeding	Age at diagnosis (d)	DOL predisease sample	DOL disease sample	DOL postdisease sample
NEC	180	23	500	Vaginal	Breast milk	16	3	18	33
	161	25	700	Vaginal	Breast milk	31	14	29	46
	199	25	725	Vaginal	Breast milk	25	10	23	41
	171	26	790	Vaginal	Breast milk	19	4	19	35
	174	29	1,350	Cesarean	Formula	26	5	26	41
	139	30	1,470	Cesarean	Breast milk	28	16	28	31
LOS	181	23	570	Vaginal	Breast milk	26	8	25	40
	130	27	1,000	Cesarean	Breast milk	20	7	20	35
	166	29	1,250	Vaginal	Breast milk	10	2	10	35
	172	28	1,060	Vaginal	Breast milk	15	2	15	31
Control	143	24	620	Vaginal	Breast milk	NA		12	
	222	24	620	Vaginal	Breast milk	NA		26	
	223	24	885	Vaginal	Breast milk	NA		21	
	152	25	800	Vaginal	Breast milk	NA		30	
	186	26	840	Cesarean	Breast milk	NA		23	
	167	27	1,290	Cesarean	Formula	NA		17	
	206	28	1,255	Vaginal	Breast milk	NA		13	
	208	29	950	Cesarean	Breast milk	NA		25	
	207	29	1,580	Cesarean	Breast milk	NA		10	

LOS, late-onset sepsis; NA, not available; NEC, necrotising enterocolitis.

### Data Analysis

Proteomic data were analyzed using Progenesis LC-MS (version 4.1; Nonlinear Dynamics, Newcastle upon Tyne, UK) detecting 447 unique proteins in total. Proteins were identified from Progenesis LCMS-generated.mgf files by searching the human proteome (uniprot 2014\_10,888,808 sequences) using Mascot version 2.4 (Matrixscience, London, UK), trypsin as the digestion enzyme, Cys-carbamidomethyl as a fixed modification, and Met-oxidation as a variable modification, allowing for one missed cleavage. Peptide tolerance was set to 10 ppm and fragment mass tolerance to 10 mmu. Search results were exported from mascot in.xml format and imported back into Progenesis LC-MS for downstream analysis.

Metabolomics data were analyzed using Progenesis QI (version 1.0; Nonlinear Dynamics). Positive and negative data were processed individually and combined prior to downstream analysis. Adducts contributing >1% relative abundance based on prevalence in the pooled samples were selected for the full analysis. Features detected after 23 min were ignored, and all features with the highest mean in the blank samples were excluded from the analysis.

For all analyses, Progenesis normalization was applied where the scaling factor is the anti-log of the mean of the log (ratios). *P* values were generated in Progenesis using ANOVA, where significant features were determined based on a probability (*P*) of <0.05. Features associated with disease diagnosis were determined in accordance with **Figure 1**: firstly, features that had a significantly higher mean at disease diagnosis compared to controls (phase I) and secondly features which increased significantly at diagnosis and decreased thereafter (phase II).

MS1 profiles underwent partial least squared discriminatory analysis using SIMCA 13.0 (Umetrics, Umeå, Sweden) (23). To check that data were adhering to multivariate normalities, Hotelling's  $T^2$  tolerance limits were calculated and set at 0.95. The R2 and Q2 values for partial least squared discriminatory analysis were used to determine how reproducible and predictive the data are, respectively, where values of >0.5 represent a robust model.

### SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at <http://www.nature.com/pr>

### ACKNOWLEDGMENTS

We thank Susan McLellen (medical laboratory assistant, Newcastle upon Tyne hospitals NHS Foundation Trust) for assistance in storage of serum samples. The content is solely the responsibility of the authors.

### STATEMENT OF FINANCIAL SUPPORT

This work was supported by funding from Tiny Lives Charity (Newcastle upon Tyne, UK), Newcastle upon Tyne Hospitals NHS Charity, and in part by an unrestricted educational grant from Nestle UK.

Disclosure: The authors declare no conflicts of interest.

### REFERENCES

- Berrington JE, Stewart CJ, Embleton ND, Cummings SP. Gut microbiota in preterm infants: assessment and relevance to health and disease. *Arch Dis Child Fetal Neonatal Ed* 2013;98:F286–90.
- Stewart CJ, Nelson A, Scribbins D, et al. Bacterial and fungal viability in the preterm gut: NEC and sepsis. *Arch Dis Child Fetal Neonatal Ed* 2013;98:F298–303.
- Claud EC, Keegan KP, Brulc JM, et al. Bacterial community structure and functional contributions to emergence of health or necrotizing enterocolitis in preterm infants. *Microbiome* 2013;1:20.
- Stewart CJ, Marrs EC, Magorrian S, et al. The preterm gut microbiota: changes associated with necrotizing enterocolitis and infection. *Acta Paediatr* 2012;101:1121–7.
- Gordon P, Christensen R, Weitkamp JH, Maheshwari A. Mapping the new world of necrotizing enterocolitis (NEC): review and opinion. *EJ Neonatol Res* 2012;2:145–72.
- Embleton ND, Turnbull E, Turner S, Berrington JE. Successful blood salvaging from preterm infants: maximizing opportunities, minimizing interventions. *Acta Paediatr* 2013;102:e527–9.
- Eras Z, Oğuz S, Dizdar EA, Sari FN, Dilmen U. Serum amyloid-A levels in neonatal necrotizing enterocolitis. *J Clin Lab Anal* 2011;25:233–7.
- Ng PC, Ang IL, Chiu RW, et al. Host-response biomarkers for diagnosis of late-onset septicemia and necrotizing enterocolitis in preterm infants. *J Clin Invest* 2010;120:2989–3000.

9. Pourcyrous M, Korones SB, Yang W, Boulden TF, Bada HS. C-reactive protein in the diagnosis, management, and prognosis of neonatal necrotizing enterocolitis. *Pediatrics* 2005;116:1064–9.
10. Wilcock A, Begley P, Stevens A, Whatmore A, Victor S. The metabolomics of necrotising enterocolitis in preterm babies: an exploratory study. *J Matern Fetal Neonatal Med* 2015; e-pub ahead of print 26 July 2015.
11. Jiang P, Siggers JL, Ngai HH, Sit WH, Sangild PT, Wan JM. The small intestine proteome is changed in preterm pigs developing necrotizing enterocolitis in response to formula feeding. *J Nutr* 2008;138:1895–901.
12. Kanakoudi F, Drossou V, Tzimouli V, et al. Serum concentrations of 10 acute-phase proteins in healthy term and preterm infants from birth to age 6 months. *Clin Chem* 1995;41:605–8.
13. Göbel U, Petrich C, von Voss H, Schmelzer L. alpha1-Antitrypsin and alpha2-macroglobulin in newborn infants. I. The influence of perinatal complications. *Eur J Pediatr* 1978;129:117–24.
14. Blair JJ, Carachi R, Gupta R, Sim FG, McAllister EJ, Weston R. Plasma alpha fetoprotein reference ranges in infancy: effect of prematurity. *Arch Dis Child* 1987;62:362–9.
15. Bellini C, Bonacci W, Parodi E, Serra G. Serum alpha-fetoprotein in newborns. *Clin Chem* 1998;44:2548–50.
16. Lerman J, Strong HA, LeDez KM, Swartz J, Rieder MJ, Burrows FA. Effects of age on the serum concentration of alpha 1-acid glycoprotein and the binding of lidocaine in pediatric patients. *Clin Pharmacol Ther* 1989;46:219–25.
17. Morrow AL, Lagomarcino AJ, Schibler KR, et al. Early microbial and metabolomic signatures predict later onset of necrotizing enterocolitis in preterm infants. *Microbiome* 2013;1:13.
18. Cetinkaya M, Ozkan H, Köksal N, Akaci O, Özgür T. Comparison of the efficacy of serum amyloid A, C-reactive protein, and procalcitonin in the diagnosis and follow-up of necrotizing enterocolitis in premature infants. *J Pediatr Surg* 2011;46:1482–9.
19. Nanthakumar N, Meng D, Goldstein AM, et al. The mechanism of excessive intestinal inflammation in necrotizing enterocolitis: an immature innate immune response. *PLoS One* 2011;6:e17776.
20. Claud EC. Neonatal necrotizing enterocolitis -inflammation and intestinal immaturity. *Antiinflamm Antiallergy Agents Med Chem* 2009;8:248–59.
21. Stewart CJ, Marrs EC, Nelson A, et al. Development of the preterm gut microbiome in twins at risk of necrotising enterocolitis and sepsis. *PLoS One* 2013;8:e73465.
22. Proc JL, Kuzyk MA, Hardie DB, et al. A quantitative study of the effects of chaotropic agents, surfactants, and solvents on the digestion efficiency of human plasma proteins by trypsin. *J Proteome Res* 2010;9:5422–37.
23. Eriksson L, Johansson E, Kettaneh-Wold N. Multi-and Megavariate Data Analysis, Part 2, Advanced Applications and Method Extensions. Sweden: MKS Umetrics AB, 2006.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>