












# Genome-wide postnatal changes in immunity following fetal inflammatory response

Daniel Costa<sup>1,2</sup> , Núria Bonet<sup>3</sup> , Amanda Solé<sup>2,4,5</sup> , José Manuel González de Aledo-Castillo<sup>6</sup> , Eduard Sabidó<sup>2,4,5</sup> , Ferran Casals<sup>3</sup> , Carlota Rovira<sup>7</sup> , Alfons Nadal<sup>8</sup> , Jose Luis Marin<sup>9</sup> , Teresa Cobo<sup>9</sup>  and Robert Castelo<sup>2,10</sup> 

1 Department of Pediatrics, Hospital de Figueres, Spain

2 Department of Experimental and Health Sciences, Universitat Pompeu Fabra (UPF), Barcelona, Spain

3 Genomics Core Facility, Department of Experimental and Health Sciences, Universitat Pompeu Fabra (UPF), Barcelona, Spain

4 Proteomics Unit, Centre de Regulació Genòmica (CRG), Barcelona, Spain

5 Barcelona Institute of Science and Technology (BIST), Barcelona, Spain

6 Inborn Errors of Metabolism Section, Laboratory of Biochemistry and Molecular Genetics, Hospital Clínic, Barcelona, Spain

7 Hospital Sant Joan de Déu, Barcelona, Spain

8 Department of Pathology, Hospital Clínic, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), University of Barcelona, Spain

9 Hospital Clínic, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Centre for Biomedical Research on Rare Diseases (CIBER-ER), University of Barcelona, Spain

10 Research Programme on Biomedical Informatics, Institut Hospital del Mar d'Investigacions Mèdiques (IMIM), Barcelona, Spain

## Keywords

fetal inflammatory response; postnatal immunity; preterm birth; proteomics; transcriptomics

## Correspondence

R. Castelo, Department of Experimental and Health Sciences, Universitat Pompeu Fabra (UPF), Barcelona, Spain; Research Programme on Biomedical Informatics, Institut Hospital del Mar d'Investigacions Mèdiques (IMIM), Barcelona, Spain  
Tel: +34933160514

E-mail: robert.castelo@upf.edu  
and

T. Cobo, Hospital Clínic, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Centre for Biomedical Research on Rare Diseases (CIBER-ER), University of Barcelona, Barcelona, Spain  
Tel: +34932279931

E-mail: tcobo@clinic.cat

The fetal inflammatory response (FIR) increases the risk of perinatal brain injury, particularly in extremely low gestational age newborns (ELGANs, < 28 weeks of gestation). One of the mechanisms contributing to such a risk is a postnatal intermittent or sustained systemic inflammation (ISSI) following FIR. The link between prenatal and postnatal systemic inflammation is supported by the presence of well-established inflammatory biomarkers in the umbilical cord and peripheral blood. However, the extent of molecular changes contributing to this association is unknown. Using RNA sequencing and mass spectrometry proteomics, we profiled the transcriptome and proteome of archived neonatal dried blood spot (DBS) specimens from 21 ELGANs. Comparing FIR-affected and unaffected ELGANs, we identified 782 gene and 27 protein expression changes of 50% magnitude or more, and an experiment-wide significance level below 5% false discovery rate. These expression changes confirm the robust postnatal activation of the innate immune system in FIR-affected ELGANs and reveal for the first time an impairment of their adaptive immunity. In turn, the altered pathways provide clues about the molecular mechanisms triggering ISSI after FIR, and the onset of perinatal brain injury.

## Databases

EGAS00001003635 (EGA); PXD011626 (PRIDE).

## Abbreviations

AF, amniotic fluid; ANC, absolute neutrophil count; DBS, dried blood spot; DE, differentially expressed; ELGAN, extremely low gestational age newborn; FDR, false discovery rate; FIR, fetal inflammatory response; GA, gestational age; GO, gene ontology; IAI, intra-amniotic infection; IQR, interquartile range; ISSI, intermittent or sustained systemic inflammation; IVH, intraventricular hemorrhage; MAS, macrophage activation syndrome; MDS, multidimensional scaling; MDSC, myeloid-derived suppressor cell; MIR, maternal inflammatory response; NEC, necrotizing enterocolitis; OR, odds ratio; PDA, patent ductus arteriosus; PRR, pattern recognition receptor; RDS, respiratory distress syndrome; ROP, retinopathy of prematurity; ROS, reactive oxygen species; TREC, T-cell receptor excision circle; UC, umbilical cord; WBC, white blood cell; WMD, white matter disease.

[Correction added on 15 December 2020, after first online publication: Peer review history is not available for this article, so the peer review history statement has been removed.]

(Received 14 April 2020, revised 22 July 2020, accepted 28 September 2020)

doi:10.1111/febs.15578

## Introduction

Intra-amniotic infection (IAI), one of the main causes of spontaneous preterm birth, can lead to a maternal (MIR) and a fetal (FIR) inflammatory response, which can be detected in placental tissues [1] (Table 1). FIR can damage the fetus presumably by triggering a local (e.g., lungs, skin) and systemic inflammation, which can lead to further distant organ injury, such as perinatal brain damage [2,3]. Understanding the molecular dynamics of FIR and its downstream alterations before and after birth is thus crucial to developing diagnostic and therapeutic strategies that improve the clinical outcome of extreme premature infants.

In a previous work, we found that FIR triggers a broad and complex transcriptional response in umbilical cord (UC) tissue [4], which others have shown to correlate with cognitive impairment at 10 years of age [5]. The molecular and cellular alterations induced by FIR persist after birth [6,7] and interact with postnatal inflammation-initiating illnesses [8]. This interaction

ultimately leads to what is known as an intermittent or sustained systemic inflammation (ISSI) [9,10]. ISSI is strongly associated with perinatal brain injury and developmental disorders in ELGANs [10] (extremely low gestational age newborn, i.e., birth before the 28 week of gestation). The joint presence of placental inflammation and ISSI appears to increase the risk of perinatal brain injury, highlighting the important contribution of postnatal systemic inflammation to that risk [11,12]. However, little is known about the extent of postnatal molecular changes associated with FIR, which can contribute to understanding the link between FIR and ISSI and consequently provide an array of candidate biomarkers and therapeutic targets for perinatal brain injury.

Here, we analyze the transcriptome and proteome of archived dried blood spot (DBS) specimens from ELGANs and identify RNA and protein expression changes in whole blood that are significantly associated with FIR. These molecular changes provide, with an unprecedented level of resolution, a snapshot of the pathways participating in ISSI as a result of the exposure to FIR.

## Results

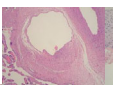
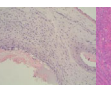
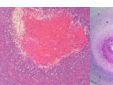
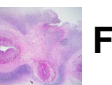
### Clinical characteristics of ELGANs and molecular profiling of DBS samples

The samples analyzed in this study were from 21 ELGANs (8 females and 13 males), who met the eligibility criteria (see [Methods](#)). Histological acute chorioamnionitis was diagnosed in 15 cases. MIR was in stage 1 in two cases, stage 2 in one case, and stage 3 in 12 cases. Ten of the 12 cases with stage 3 MIR showed FIR: four in stage 1, five stage 2, and one stage 3 (Fig. 1). Other clinical characteristics and outcomes of ELGANs are presented in Table 2, separately for those with ( $n = 10$ ) and without FIR ( $n = 11$ ).

Following the standard protocols at the hospital, all analyzed infants received antibiotic treatment. All but

**Table 1.** Staging of the maternal (MIR) and fetal (FIR) inflammatory responses in ascending intrauterine infection (Khong *et al.* [1], p. 709).

Stage	MIR	FIR
1	Acute subchorionitis or chorionitis	Chorionic vasculitis or umbilical phlebitis
2	Acute chorioamnionitis: polymorphonuclear leukocytes extend into fibrous chorion and/or amnion	Involvement of the umbilical vein and one or more umbilical arteries
3	Necrotizing chorioamnionitis: karyorrhexis of polymorphonuclear leukocytes, amniocyte necrosis, and/or amnion basement membrane hypereosinophilia	Necrotizing funisitis

		Control	1	2	3	
	Stage					FIR
3	MIR	2	4	5	1	
2	MIR	1	0	0	0	
1	MIR	2	0	0	0	
Control	MIR	6	0	0	0	

**Fig. 1.** Pathological examination of the placenta. Number of neonates cross-classified by the diagnosis of the stage of the maternal (MIR) and the fetal (FIR) inflammatory response. Original magnifications: MIR Control 40 $\times$ ; MIR Stage 1 20 $\times$ ; MIR Stage 2 40 $\times$ ; MIR Stage 3 20 $\times$ ; FIR Control 10 $\times$ ; FIR Stage 1 20 $\times$ ; FIR Stage 2 40 $\times$ ; FIR Stage 3 2 $\times$ .

three mothers of FIR-affected ELGANs received at least two doses of a complete course of antenatal steroids, which in our study is compatible with the null hypothesis of no association with FIR (Table 2). None of the infants of the cohort received postnatal treatment with glucocorticoids. FIR-affected ELGANs presented a higher frequency of microbial invasion of the amniotic cavity (89% vs 29%, two-tailed Fisher's exact test  $P = 0.035$ ), higher levels of interleukin-6 (IL-6) in amniotic fluid (AF) ( $t$ -test  $P = 0.034$ ) and lower levels of AF glucose ( $t$ -test  $P = 0.018$ ), than unaffected ones. The maternal blood C-reactive protein (CRP) levels at admission were also significantly higher in FIR-affected ELGANs ( $t$ -test  $P = 0.041$ ). Our data are compatible with no differences in neonatal morbidities between FIR-affected and unaffected ELGANs, although a nonsignificant higher frequency of neonatal sepsis and intraventricular hemorrhage was observed in FIR cases. Mean gestational age (GA) and birth weight were similar between the two groups, averaging 26 weeks and 860 g, respectively. Such an extremely low birth weight, with a lowest value of 580 g and up to 8 neonates under 750 g, precludes a standard blood extraction after birth for the purpose of molecular profiling in research. For this reason, DBS samples constitute a minimally invasive option to profile the transcriptome and proteome in peripheral blood from ELGANs.

Total RNA sequencing of the 21 archived DBSs produced more than 500 million paired-end reads, which we transformed into a table of 25 221 gene-level

**Table 2.** Clinical characteristics of ELGANs. Baseline maternal characteristics, biomarkers of intra-amniotic infection and inflammation and perinatal outcomes. ALC, absolute lymphocyte count; MIAC, microbial invasion of the amniotic cavity defined by a positive AF culture. Intraventricular hemorrhage: any grade between I and IV according to the Papile *et al.* [13] grading system. Sepsis: confirmed or suspected bacterial blood infection, which is classified as early if within 72 h after birth and classified as late if after the third postnatal day. Categorical characteristics were tested with respect to FIR status using a two-tailed Fisher's exact test. Continuous characteristics were tested with respect to FIR status using a  $t$ -test.

Clinical characteristics	No ( $n = 11$ )	FIR FIR ( $n = 10$ )	$P$ value
Baseline maternal characteristics			
Nulliparity: $n$ , (%)	9 (81.8)	5 (50.0)	0.183
Preterm labor: $n$ , (%)	3 (27.3)	3 (30.0)	1.000
Preterm, pre-labor rupture of membranes: $n$ , (%)	1 (9.1)	4 (40.0)	0.149
Cervical insufficiency: $n$ , (%)	2 (18.2)	3 (30.0)	0.635
Placenta abruption: $n$ , (%)	4 (36.4)	0 (0.0)	0.090
Preeclampsia: $n$ , (%)	1 (9.1)	0 (0.0)	1.000
Complete course (at least 2 doses) of antenatal steroids: $n$ , (%)	11 (100)	7 (70.0)	0.090
Maternal CRP concentration at admission: (mg-dL <sup>-1</sup> ), mean $\pm$ SD	2.3 $\pm$ 3.4	5.8 $\pm$ 3.8	0.041
Biomarkers of intra-amniotic infection and inflammation			
MIAC: $n$ , (%)	2 (28.6)	8 (88.9)	0.035
Concentration of IL-6 in AF: (pg-mL <sup>-1</sup> ), (Q1, Q3)	21 153.5 (6695.50, 35 975.75)	112 271.5 (34 968.07, 162 866.25)	0.034
Concentration of glucose: (mg-dL <sup>-1</sup> ), (Q1, Q3)	23.6 (10, 37)	2.4 (0, 2)	0.018
Clinical chorioamnionitis: $n$ , (%)	0 (0.0)	3 (30.0)	0.090
Histological chorioamnionitis: $n$ , (%)	5 (45.5)	10 (100.0)	0.012
Perinatal outcomes			
Cesarean section: $n$ , (%)	8 (72.7)	3 (30.0)	0.086
Male: $n$ , (%)	7 (63.6)	6 (60.0)	1.000

**Table 2.** (Continued).

Clinical characteristics	No ( <i>n</i> = 11)	FIR FIR ( <i>n</i> = 10)	<i>P</i> value
GA: (week), mean ± SD	26.5 ± 1.3	26.2 ± 1.2	0.489
Birth weight: (g), mean ± SD	841 ± 138	880 ± 159	0.552
Respiratory distress syndrome: <i>n</i> , (%)	5 (45.5)	5 (50.0)	1.000
Patent ductus arteriosus: <i>n</i> , (%)	6 (54.5)	6 (60.0)	1.000
Intraventricular hemorrhage: <i>n</i> , (%)	1 (9.1)	5 (50.0)	0.063
White matter disease: <i>n</i> , (%)	1 (9.1)	2 (20.0)	0.586
Retinopathy of prematurity: <i>n</i> , (%)	5 (45.5)	3 (37.5) 2 missing obs.	1.000
Necrotizing enterocolitis: <i>n</i> , (%)	1 (9.1)	0 (0.0)	1.000
Sepsis: <i>n</i> , (%)	1 (9.1)	5 (50.0)	0.063
DBS heel sampling days > 3: <i>n</i> , (%)	4 (36.4)	5 (50.0)	0.678
WBC count at birth: (10 <sup>9</sup> ·L <sup>-1</sup> ), mean ± SD	8.6 ± 3.5	23.0 ± 15.6	0.017
Postnatal WBC count within first 8 postnatal days: (10 <sup>9</sup> ·L <sup>-1</sup> ), mean ± SD	13.2 ± 5.2	44.1 ± 13.4	< 0.001
Maximum ANC within first 8 postnatal days: (10 <sup>9</sup> ·L <sup>-1</sup> ), mean ± SD	6.2 ± 4.3	29.0 ± 10.8	< 0.001
ALC at birth: (10 <sup>9</sup> ·L <sup>-1</sup> ), mean ± SD	4.3 ± 1.2	6.7 ± 3.1	0.039
Lymphocyte percentage at birth: (%), mean ± SD	54.3 ± 14.5	35.4 ± 16.3	0.012
Maximum ALC within first 8 postnatal days: (10 <sup>9</sup> ·L <sup>-1</sup> ), mean ± SD	6.0 ± 3.2	10.8 ± 4.1	0.009
Maximum lymphocyte percentage within first 8 postnatal days: (%), mean ± SD	46.9 ± 17.1	25.4 ± 10.4	0.003
Maximum CRP within first 8 postnatal days: (md·dL <sup>-1</sup> ), mean ± SD	1.0 ± 2.2	3.5 ± 5.3	0.187

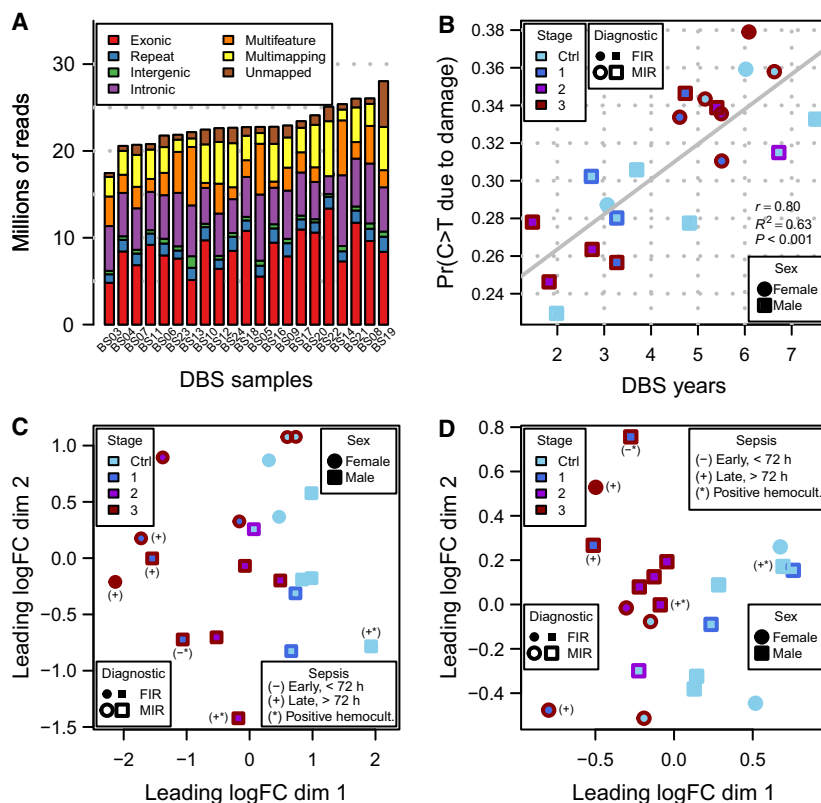
summarized count expression profiles. Nearly half of these genes (11 279) showed reliable levels of expression in at least 6 of the 21 samples and were those selected for further analysis (see [Methods](#)). [Figure 2A](#)

shows the distribution of sequencing depth across different genomic elements for each sample, where only about  $\frac{1}{3}$  of the total sequencing depth derives from exonic reads. The small fraction of reads mapping to intergenic regions suggests no DNA contamination. DBS samples in this cohort were stored at room temperature for 1–7 years until RNA extraction and sequencing. Over this period of time, degradation and fragmentation processes affect RNA integrity and we thus investigated whether, as observed in ancient DNA samples [14], there is a detectable damage to the RNA due to cytosine deamination on such a short timescale (see [Methods](#)). Indeed, we found a significant linear relationship ( $r = 0.8$ ,  $P < 0.001$ ) between the probability of a cytosine to thymine (C > T) substitution due to damage and the elapsed time between DBS extraction and RNA sequencing ([Fig. 2B](#)). Additionally, mass spectrometry proteomics on DBSs produced 649 quantified protein expression profiles ([Data S1](#)), from which we selected 245 as being reliably expressed in at least 6 of the 21 samples (see [Methods](#)).

Sample-level changes of transcriptomics data, projected in two dimensions ([Fig. 2C](#)), show that gene-level expression profiles derived from exonic reads capture the differences in placental inflammation throughout the different stages of MIR and FIR. Analogously, DBS protein expression profiles also show a clear separation between infants with and without MIR and FIR ([Fig. 2D](#)). In summary, transcriptomics and proteomics of archived DBS samples can produce a large catalog of molecular measurements that capture clinical features of the perinatal period.

### Neonatal DBSs contain gene and protein expression changes associated with FIR

The differential expression analysis of RNA-seq transcriptomic profiles between FIR- and non-FIR-exposed ELGANs identified 782 differentially expressed (DE) genes under an experiment-wide significance threshold of 5% false discovery rate (FDR) and a minimum 50%-fold change ([Fig. 3A](#), [Table S1](#)). We replicated similar magnitude changes from DE genes on a logarithmic scale ( $r = 0.75$ ,  $R^2 = 0.56$ ,  $P < 0.001$ ) in the analogous differential expression analysis of a DBS gene expression dataset [15] from the US ELGAN cohort ([Fig. 3B](#), [Table S2](#) and [Figs 4–7](#)), despite differences in sample preparation and gene expression profiling technology (see [Methods](#)). The hierarchical clustering of the 21 transcriptomic samples, using the 782 DE genes, leads to a perfect separation of the neonates who were affected by FIR, from those that were unaffected ([Fig. 8](#)).



**Fig. 2.** Transcriptomics and proteomics profiling of archived DBSs. (A) Sequencing depth per sample, broken by origin of sequence reads. (B) Probability of a C > T deamination due to damage, as function of DBS age in years. (C) Sample differences in terms of fold changes between transcriptomic samples in base-2 logarithmic scale (logFC), projected along the x- and y-axes. Distance between points is proportional to the dissimilarity between samples. (D) Same as (C) but calculated from proteomic samples.

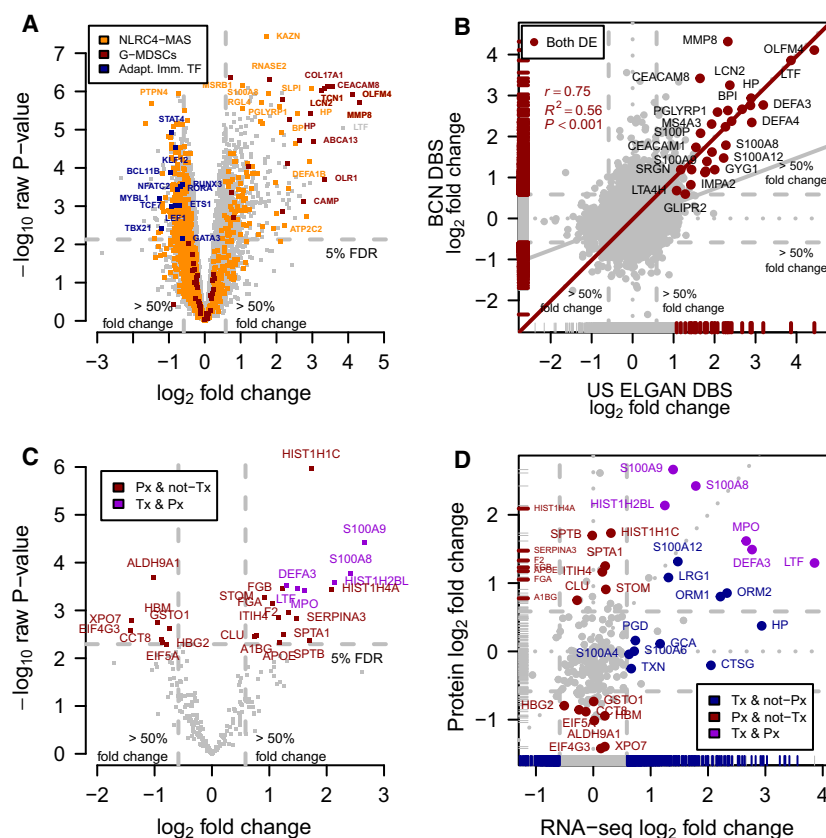
A further question is whether there are RNA changes in peripheral blood across the different stages of MIR and FIR. Indeed, we found 58 genes with FDR < 5%, for which a combination of MIR and FIR stages, reflecting an increasing exposure to prenatal inflammation, led to a minimum increase of 50% in gene expression levels for each additional combined inflammatory stage (Fig. 9, Table S3).

The differential expression analysis of proteomic profiles between FIR-exposed neonates and those who were unexposed identified 27 DE proteins at 5% FDR with a minimum 50%-fold change (Fig. 3C, Table S4). Figure 3D shows a comparison between log<sub>2</sub>-fold changes of the transcriptomics and proteomics analyses. Six of the 27 DE proteins are encoded by genes that are also DE by RNA-seq, while for 14 of them, their corresponding genes do not show fold changes above 50% in the transcriptomics data. The remaining seven DE proteins do not have minimum levels of expression in the RNA-seq assay. Among them, F2, FGA, FGB, and A1BG are secreted from the liver into the blood, and this may also be the case for SERPINA3 and APOE, which have a biased expression toward the liver. Overall, these results show that in peripheral blood from ELGANs there are sizable numbers of transcriptomic and proteomic changes associated with FIR that can be detected in archived DBS.

### FIR is associated with a postnatal activation of the innate immune system

Blood protein biomarkers have helped in the discovery of a postnatal systemic inflammatory hit following FIR [6]. However, the extent of such molecular changes was previously unknown. Our transcriptomic data from DBS samples show the postnatal upregulation in FIR of many genes involved in the innate immune response (Fig. 3A; Table S1). Such genes include those coding for pattern recognition receptors (PRRs, e.g., *NLRC4*, *CLEC4E*, *FPRI*), inflammatory transcription factors (e.g., *CEBPD*, *KLF5*), acute phase response proteins (*HP*, *ORM1*), as well as multiple inflammatory mediators including cytokines (e.g., *IL18*, *IL1R1*, *RETN*), calgranulins (*S100A8*, *S100A9*, *S100A12*), chemokines (e.g., *CXCR1*, *CXCL1*), complement (e.g., *C5AR2*, *C3AR1*), lipid mediators (e.g., *LTA4H*, *ALOX5AP*, *CYP4F3*), proteolytic enzymes (e.g., *MMP8*, *MMP9*) and adhesion molecules (e.g., *ICAM3*, *ITGAM*). Our proteomic data from DBS samples (Fig. 3C, Table S4) also show protein overexpression in FIR of diverse inflammatory mediators of the innate immune system, such as the S100A8, S100A9, LTF, MPO and DEFA3 proteins. A functional enrichment analysis using the Gene Ontology (GO) database (Fig. 8) shows GO terms associated





**Fig. 3.** Differential expression analysis between FIR-affected and unaffected ELGANs, of transcriptomic and proteomic profiles from archived DBSs. (A) Volcano plot of transcriptomics data. Raw  $P$  values in negative logarithmic scale on the y-axis as a function of the  $\log_2$ -fold change on the x-axis. Orange dots highlight genes from a NLR4-MAS signature [16], red dots highlight genes from a G-MDSCs signature [17], and blue dots highlight transcription factor coding genes involved in regulating adaptive immunity. (B) Fold changes of this study on the y-axis as a function of fold changes in data from the US ELGAN cohort in the x-axis. (C) Volcano plot of proteomics data. (D) Fold changes from proteomics data on the y-axis as a function of fold changes from transcriptomics data on the x-axis.

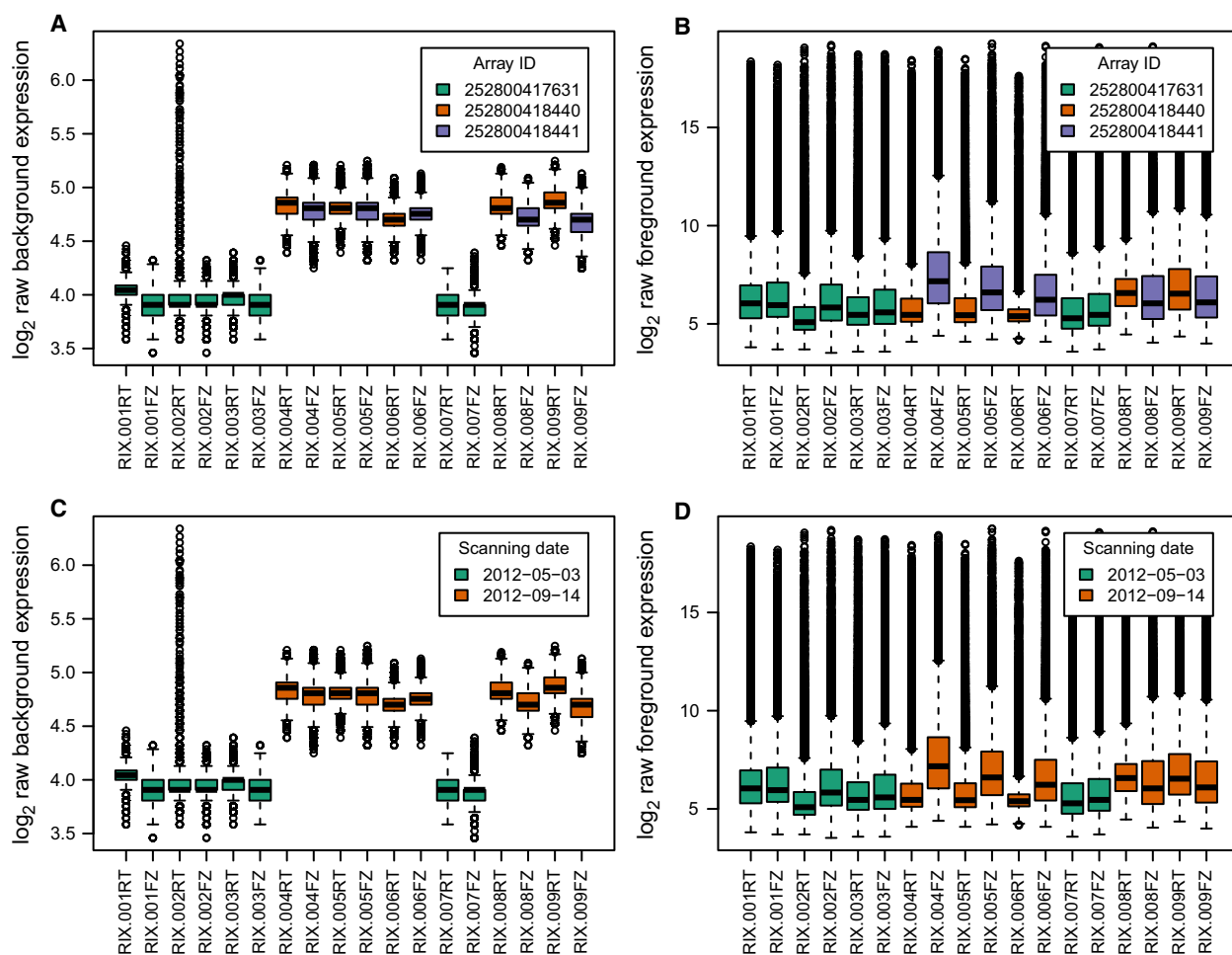
with the innate immune system, such as *innate immune response in mucosa* (OR = 17.7) and *defense response to fungus* (OR = 14.3), which were significantly enriched (FDR < 10%) in upregulated genes (Table S5; see Methods). We also observed a significant enrichment of DE genes in curated gene sets of the innate immune response, such as the InnateDB [18] database (one-tailed Fisher's exact  $P < 2.2 \times 10^{-16}$ , OR = 3.2).

Moreover, we found that FIR upregulates the expression of the NLR4 inflammasome and IL-1 family-related protein-coding genes (e.g., *NLR4*, *PYCARD*, *IL18*, *IL1R1*), including a significant overlap (Fig. 3A, one-tailed Fisher's exact  $P < 2.2 \times 10^{-16}$ , OR = 10.4) of a transcriptional signature of the NLR4-infantile-onset macrophage activation syndrome (MAS) [16] with our set of 782 DE genes. Consistent with these results, we observed the upregulation of numerous genes involved in neutrophil biology and function (Table S6), significantly enriching (FDR < 10%, OR > 1.5) GO terms related to neutrophil activation [19], such as *neutrophil migration* (OR = 6.3) or *positive regulation of neutrophil chemotaxis* (OR = 7.5); see Fig. 8 and Table S5.

Previous studies of postnatal molecular changes associated with FIR in ELGANs [6,7,10] were based on a handful of gene and protein biomarkers. The gene and protein expression changes described here constitute, to our knowledge, the most comprehensive description to date of the postnatal activation of the innate immune system in ELGANs affected by FIR. Our data suggest that FIR leads to a postnatal robust activation of the IL-18/IL-1 axis through the NLR4-inflammasome that can induce neutrophil activation, contributing to a systemic inflammation.

### FIR is associated with the impairment of adaptive immunity and thymic function through the expansion of granulocytic myeloid-derived suppressor cells

The mechanism by which antenatal FIR leads to a postnatal systemic inflammation is unknown. Among several plausible alternatives, Dammann and Leviton [9] suggested an alteration in the number of T cells and their regulation. A closer look at the hierarchical clustering of genes in Fig. 8 reveals an enrichment of downregulated transcription factor coding genes (one-

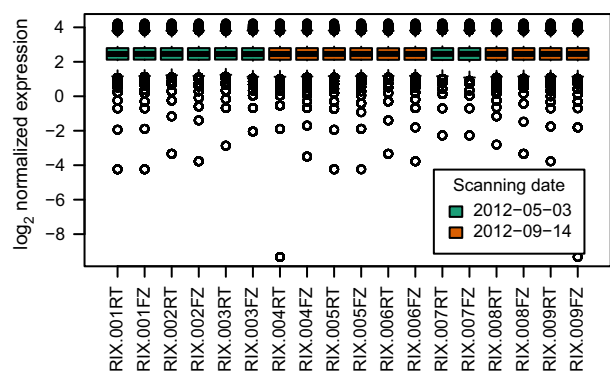


**Fig. 4.** Distribution of raw expression values from the US ELGAN cohort microarray data. (A, C) Expression values from background probes. (B, D) Expression values from foreground probes. (A, B) Samples colored by array identifier. (C, D) Samples colored by scanning date. In the boxplots, the horizontal solid line indicates the median, the box limits correspond to the 25th and 75th percentile that span the IQR, while the whiskers, or error bars, extend to 1.5 times the IQR away from the median and values exceeding that distance, if any, are depicted by open circles.

tailed Fisher's exact test  $P < 0.001$ ,  $OR = 2.9$ , Table S7) in FIR-affected ELGANs. Many of these genes are regulators of T-cell development and activation, such as *TBX21*, *GATA3* and *RORA*, respectively involved in T-helper(h) fate of CD4+ T cells in Th1, Th2 and Th17 lineages (Fig. 3A). More generally, we found the downregulation of key genes associated with the IL-7 and IL-2 signaling pathways (e.g., *IL7R*, *BCL2*, *PIK3AP1*, *IL2RB*, *CD3E*, *CD3G*) involved in T-cell development (Tables S1 and S8). The T-cell receptor signaling pathway (e.g., *CD3E*, *CD3G*, *LCK*, *FYN*, *ZAP70*, *CD247*) and T-cell differentiation co-stimulatory receptor genes (e.g., *CD28*, *ICOS*, *SLAMF1*) were also downregulated [20]. Likewise, the inhibitory and stimulatory genes *KLRC1*, *KLRC3*, *KLRC4* that belong to the natural killer (NK) cell lectin-like receptors of the NK gene complex, and genes

involved in NK cell cytotoxicity, such as *PRF1*, *SH2D1B* genes, were also downregulated in FIR-affected ELGANs [20]. Interestingly, we observed a significant overlap of downregulated genes with an NLRC4-infantile-onset MAS signature (Fig. 3A, one-tailed Fisher's exact  $P < 2.2 \times 10^{-16}$ ,  $OR = 21.5$ ) [16], which is associated with cytotoxic and regulatory T-cell dysfunction.

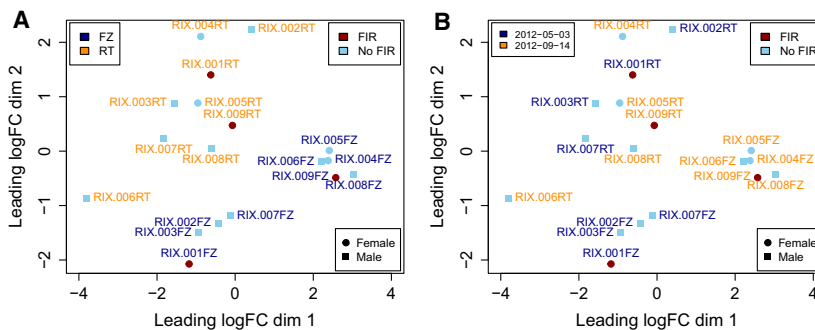
Recent thymic emigrants consist of phenotypically and functionally immature T cells that have recently egressed from the thymus, which predominate in the fetus and newborns. These cells can be identified by their content of T-cell receptor excision circles (TRECs) [21], which are DNA fragments excised during the rearrangement of segments of T-cell receptor genes in the thymus. TRECs are used as a measure of the number of T lymphocytes in newborns and,



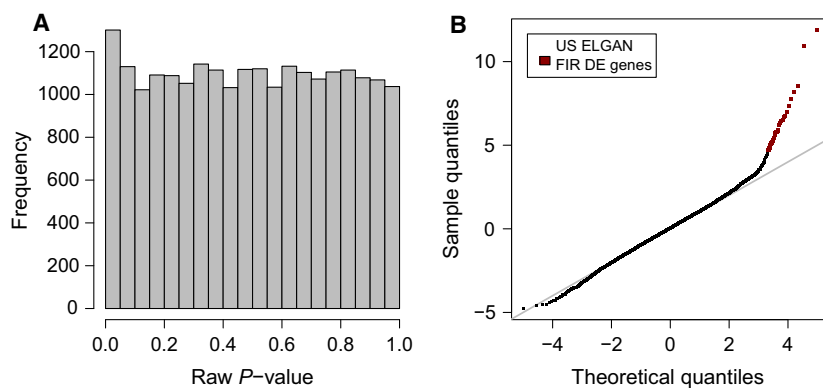
**Fig. 5.** Distribution of background-corrected and quantile-normalized expression values from the US ELGAN cohort microarray data. In the boxplots, the horizontal solid line indicates the median, the box limits correspond to the 25th and 75th percentile that span the IQR, while the whiskers, or error bars, extend to 1.5 times the IQR away from the median and values exceeding that distance, if any, are depicted by open circles.

indirectly, as an indicator of a possible thymus dysfunction. We found downregulation of the genes coding for the PTK7 and S1PR1 proteins in FIR-affected ELGANs. PTK7 is a cell marker of CD4<sup>+</sup> T-cell recent thymic emigrants used for assessing thymic activity [22], while the S1PR1 receptor expression in CD4<sup>+</sup> T cells is required for their survival and thymic emigration [23]. Consistent with these results, FIR-affected ELGANs had significantly lower levels of TRECs in the analyzed DBSs ( $P = 0.0432$ ; Fig. 10A, see Methods), a significantly higher postnatal neutrophil to lymphocyte ratio ( $P = 0.002$ ; Fig. 10B), and a significantly lower percent of all leukocytes that are lymphocytes at birth ( $P = 0.012$ ; Table 2) and during the first postnatal week [24] ( $P = 0.003$ ; Fig. 10C).

Myeloid-derived suppressor cells (MDSCs) are immature myeloid cells that can develop under certain pathological conditions and are central regulators of immune responses. The hallmark of MDSCs is their immune suppressive activity, and according to their

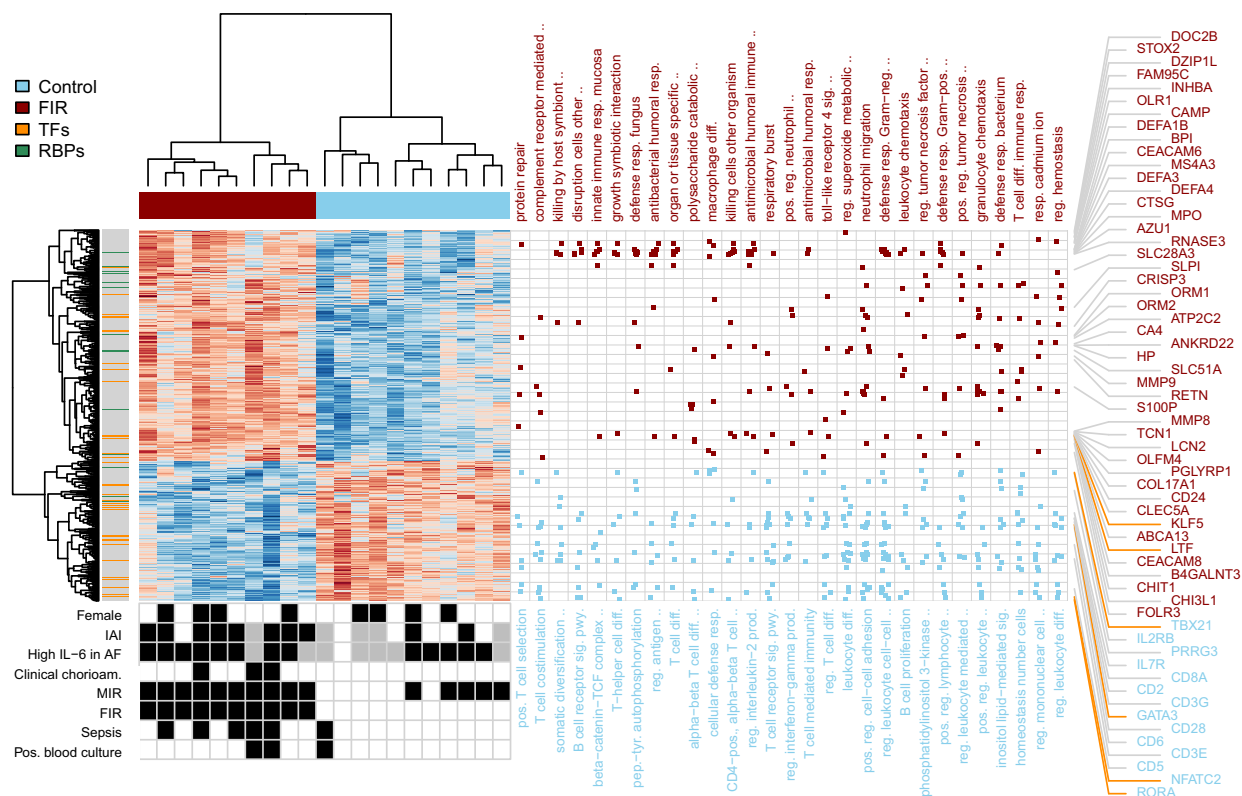


**Fig. 6.** Sample differences from the US ELGAN cohort microarray data, in terms of fold changes between samples in base-2 logarithmic scale. (A) Sample labels are colored by storage condition (FZ: frozen; RT: room temperature). (B) Sample labels are colored by microarray scanning date. There is a clear effect of storage condition, an effect of scanning date mostly in frozen samples and no apparent sex or FIR effect.



**Fig. 7.** Differential expression between FIR and non-FIR samples from the US ELGAN cohort microarray data. (A) Distribution of raw  $P$ -values (B) Quantile–quantile plot of moderated  $t$ -statistics. Red dots indicate differentially expressed genes with FDR < 5% and minimum 1.5-fold change.



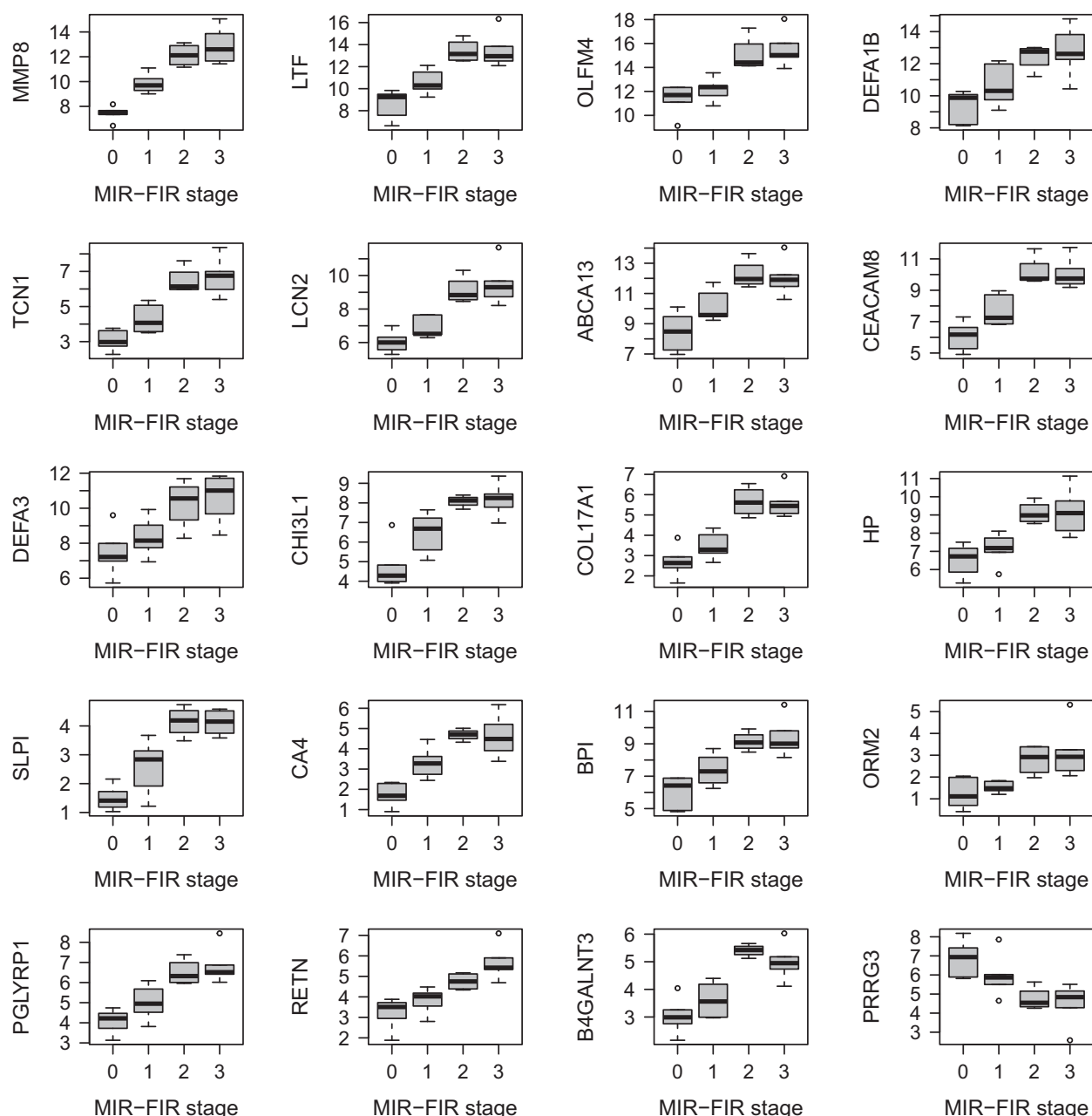


**Fig. 8.** FIR gene expression signature and functional enrichment analysis. Heatmap of expression values for 782 DE genes with FDR < 5% and minimum 1.5-fold change between FIR-affected and unaffected ELGANS, obtained after removing FIR-unrelated variability. Dendrograms on the x- and y-axes represent the hierarchical clustering of samples and genes, respectively. Leaves on the gene dendrogram are color coded according to whether genes encode for transcription factors (orange) or RNA-binding proteins (green), while those on the sample dendrogram are color coded according to FIR status, where red indicates samples derived from FIR-affected infants and blue unaffected. The right dot-matrix represents DE genes (y-axis) belonging to GO terms (x-axis) significantly enriched (FDR < 10% and odds ratio -OR > 1.5) by upregulated (top) and downregulated (bottom) DE genes. The bottom-left dot-matrix represents the presence or absence of phenotypes, with gray indicating missing data.

phenotype, MDSCs can be classified into monocytic (M)-MDSCs or granulocytic(G)-MDSCs. Postnatally, there is a high level of circulating MDSCs, predominantly G-MDSCs, which correlates with inflammatory markers of perinatal infection [25,26]. Neonatal MDSCs display reduced apoptosis and immunosuppressive activity after bacterial infection [27]. Indeed, these cells are potent inhibitors of adaptive immune responses, particularly of T and NK cells, by different mechanisms, including the production of reactive oxygen species (ROS) [28–30], prostaglandin E2 (PGE2), and the depletion of L-arginine by the arginase 1 enzyme (ARG1) [31,32]. In our data, we found that the *ARG1* gene was 2.5-fold upregulated in FIR-affected ELGANS, as well as the *OLR1* gene, encoding the lectin-type oxidized LDL receptor 1 (LOX1), a

marker of G-MDSCs [17], which was over 10-fold upregulated. We also found a significant overlap (one-tailed Fisher's exact  $P < 2.2e-16$ , OR = 26.4) between FIR upregulated genes and a G-MDSC expression signature [17] (Fig. 3A).

One of the mechanisms by which G-MDSCs suppress T-cell responses is the release of ROS molecules. We found an upregulation of genes coding for proteins essential to ROS production such as *CYBB*, *NCF1*, *NCF4*, *MPO*, *RAB27A*, as well as genes coding for antioxidant enzymes (e.g., *SOD2*, *TXN* and the Msr family -*MsrA*, *MsrB1*, *MsrB2*, *MsrB3*). The Msr gene family encodes proteins of the methionine sulfoxide reductase system that reduces protein methionine oxidation, a post-translational alteration that correlates with oxidative stress [33]. In summary, our data

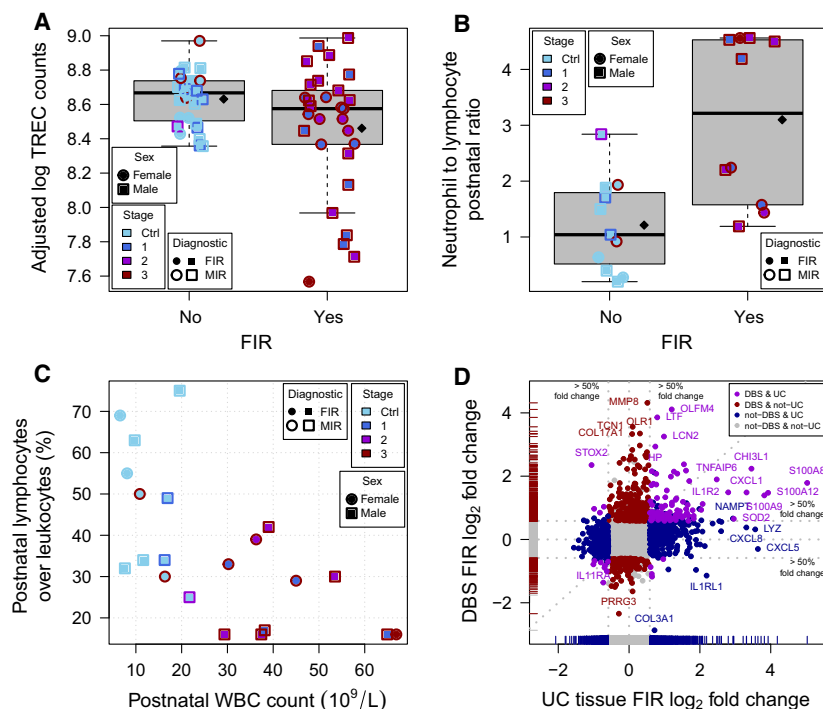


**Fig. 9.** Combined MIR and FIR linear effect on gene expression. Top-20 genes with largest fold change among DE genes with FDR < 5% and minimum 50%-fold change throughout the following combined 4 stages of MIR and FIR: 0 no MIR and no FIR; 1 MIR but no FIR; 2 MIR stage 3 and FIR stage 1; 3 MIR stage 3 and FIR stage 2 or 3. Adjusted expression values in log-CPM units are plotted in the y-axis as function of the combined stages in the x-axis. In the boxplots, the horizontal solid line indicates the median, the box limits correspond to the 25th and 75th percentile that span the IQR, while the whiskers, or error bars, extend to 1.5 times the IQR away from the median and values exceeding that distance, if any, are depicted by open circles.

suggest the expansion of G-MDSCs as a mechanism contributing to the impairment of adaptive immunity in ELGANs affected by FIR, and consequently, to a postnatal systemic inflammation sustained over time after birth.

## Discussion

We have described the largest catalog to date of postnatal blood transcriptomic and proteomic changes associated with FIR in archived DBSs from ELGANs.



**Fig. 10.** Association of FIR with clinical and molecular signatures. (A) TREC-corrected fluorescent counts on logarithmic scale by FIR status, after adjusting for sex, maximum lymphocyte count during the first postnatal week and batch. Diamonds indicate mean values. There are multiple measurements per newborn. (B) Postnatal ratio of neutrophil absolute count to absolute lymphocyte count, during the first postnatal week, by FIR status. Diamonds indicate mean values. (C) Percent of all leukocytes that are lymphocytes during the first postnatal week on the y-axis as a function of the postnatal white blood cell count on the x-axis. (D) Fold changes in this study on the y-axis as a function of FIR fold changes in UC tissue on the x-axis. In the boxplots, the horizontal solid line indicates the median, the box limits correspond to the 25th and 75th percentile that span the IQR, while the whiskers, or error bars, extend to 1.5 times the IQR away from the median and values exceeding that distance, if any, are depicted by open circles.

This catalog demonstrates that archived DBSs, collected for the newborn screening program, constitute a valuable source of genome-wide molecular information about the perinatal period. More importantly, we have shown that molecular changes in this catalog provide clues to the underlying relationships between FIR and ISSI, with common and specific alterations to each of these two conditions (Fig. 10D and Table S9).

Our data provide molecular evidence of the postnatal activation of NLRC4-inflammasome dependent mechanisms, which may contribute to the development of ISSI in FIR-affected ELGANs. After cytosolic detection of microbial and endogenous danger signals by NOD-like receptors, the NLRC4 inflammasome induces, via activation of caspase-1 (CASP1), the maturation and release of IL-1 $\beta$  and IL-18 cytokines, the bacterial clearance and the pyroptosis of immune cells. The release of interleukin-1 family cytokines such as IL-1 $\beta$  and IL-18 has potent effects on neutrophil recruitment, activation and lifespan [34]. However, the NLRC4-inflammasome hyperactivity can impair

inflammation resolution and cause diverse auto-inflammatory diseases, such as the infantile-onset MAS syndrome and enterocolitis. These disorders are characterized by a huge persistent elevation of IL-18 in the blood [16] as in neonatal sepsis [35] and, in fact, we observed a 2-fold overexpression of the *IL18* gene in FIR-affected infants. Liang *et al.* [36] reported recently the case of a 28-week preterm neonate with a *de novo* gain-of-function mutation in the *NLRC4* gene, who died of a fatal syndrome of excessive immune activation likely to have begun in utero. NLRC4 dependent mechanisms can also contribute to brain injury induced by cerebral ischemia, as has recently been described in a rodent-based model of stroke [37]. Interestingly, caffeine therapy for the apnea of prematurity, which reduces the incidence of moderate to severe neurodevelopmental disabilities in very low birth weight infants [38], can also reduce NLRC4-inflammasome activation [39].

The resolution of inflammation requires neutrophil arrest, death and removal [40]. However, a high and

persistent release of the pro-inflammatory cytokines IL-18 and IL-1 $\beta$  after NLRC4 inflammasome hyperactivation may induce the continued expansion, migration and activation of neutrophils. Moreover, molecular alterations in neonatal neutrophils, such as a reduced expression of Fas receptors, can lead to delayed neutrophil apoptosis [41,42] and the persistence of pro-inflammatory responses [43]. In FIR-affected ELGANs, we found a significantly higher neutrophil to lymphocyte ratio (Fig. 10B), the overexpression of genes involved in neutrophil biology, such as *S100A8*, *S100A9*, *IL18*, *MMP8* and *OLFM4*, and the underexpression of gasdermin B genes (Table S1). Calgranulins S100A9 and S100A8, and IL-18 can inhibit neutrophil apoptosis [44,45] and neutrophils can evade pyroptosis, which is mediated by gasdermin family proteins [46]. Furthermore, the robust production of calgranulins and the activation of the IL-1/IL-18 axis can trigger a neutrophil-associated inflammation-boosting loop binding to PRRs and triggering downstream inflammatory signaling pathways, just as in auto-inflammatory disorders [47]. Among children with septic shock, elevated *MMP8* and *OLFM4* blood gene expression levels, which are markers of neutrophil activation, are associated with higher rates of mortality and organ failure [48,49]. Similarly, other studies have also shown that preterm neonates with abnormal neuroimaging studies had increased blood neutrophil activation markers [50,51].

Chorioamnionitis can lead to an inflammation-induced immunosuppression similar to the sepsis-induced immunosuppression [52,53]. This chorioamnionitis-induced immunosuppression includes the depletion of B and T lymphocytes in spleen [54], thymic involution [55], endotoxin hyporesponsiveness [56] and lower CXCL8-producing CD4 and CD8 T cells [57]. In umbilical cord blood monocytes from very preterm newborns (GA 29–32 weeks) affected by chorioamnionitis, it has been described the downregulation of genes involved in adaptive immunity upon monocyte stimulation with *Staphylococcus epidermidis* [58], the most common cause of neonatal late-onset sepsis. The exposure of umbilical cord blood monocytes to chorioamnionitis has been also shown to alter the histone modification landscape and reduce the immune response to a second inflammatory stimulus with lipopolysaccharides [59]. The molecular events underlying this chorioamnionitis-induced immunosuppression are, however, not well characterized in extremely preterm newborns, particularly after birth. In our data, the analysis of downregulated genes provides for the first time molecular evidence of the postnatal inhibition of adaptive immune responses, particularly T

cell, in FIR-affected ELGANs (Figs 3A and 8). This evidence is consistent with the observed relationships between TRECs and FIR (Fig. 10A), as well as the lymphocyte percent among ELGANs with and without FIR and MIR (Fig. 10C). The reduction of TRECs and the downregulation of the *PTK7* and *SIPRI* genes suggest a thymic involution associated with FIR [60], which may lead to a relative increase in pro-inflammatory CD31-CD4<sup>+</sup> T cells producing TNF [61]. Perhaps that is why thymic involution has also been associated with increased risk of brain damage in ELGANs [62].

Previous findings are also consistent with our data because we found a significant overlap between FIR DE genes and a published G-MDSC signature [17] (Fig. 3A), as well as the underexpression of the *LCK* and *ZAP70* genes and the overexpression of *S100A8*, *S100A9* and *LTF* genes and protein products in FIR-affected ELGANs (Fig. 3D). In newborns, the LTF protein triggers a G-MDSCs T-cell-suppressive phenotype, mediated by the secretion of S100A8, S100A9 and PGE2 [63], which repress the expression of LCK and ZAP70 in CD4<sup>+</sup> T cells, impairing their activation [29,30,64]. Overall, our results suggest that after an initial prenatal pro-inflammatory-driven response evident in UC tissue (e.g., funisitis) and UC blood (IL-6, IL-1 $\beta$ , CXCL8), FIR is followed by an intense neonatal innate immune activation mediated by the NLRC4-inflammasome hyperactivity. As a result of these events, FIR-affected ELGANs display a robust neutrophil and G-MDSC expansion during the first postnatal days, which can lead to adaptive immune suppression, thymic involution, delayed neutrophil apoptosis and the onset of an auto-inflammatory loop, contributing to ISSI and consequently increasing the risk of perinatal brain injury.

An important feature of G-MDSCs is the production of ROS molecules, which can further lead to oxidative stress and damage. Preterm newborns have a reduced antioxidant capacity that increases the risk of ROS-induced damage, which is an important factor in numerous diseases in premature infants, such as perinatal brain damage [65]. In fact, ELGANs are born with deficient levels of selenium [66], a cofactor for the MsrB and thioredoxin enzyme activity required for proper antioxidant function. We found overexpression of genes coding for antioxidant enzymes (e.g., Msr family genes) and a 10-fold overexpression in FIR of both the *OLRI* gene, which encodes LOX1, a marker of G-MDSCs [17], and the gene coding for the heat shock protein HSPA1A. HSPA1A can activate LOX1 in UC neutrophils in response to *in vitro* exposure to microbial peptidoglycan [67], leading to the downstream production of ROS. The LTF protein, which

we found to be significantly overexpressed in both transcriptomics and proteomics data, can protect newborns from oxidative stress and damage by attenuating inflammatory responses and controlling oxidative cell injury induced by innate immune activation [68].

Using well-established biomarkers, it has been shown that antenatal and postnatal inflammation, the so-called two-inflammatory hit, increases the risk of perinatal brain damage in preterm newborns [11]. Here, we found the overexpression of multiple genes encoding proteins that can be associated with the onset of perinatal brain damage in FIR-exposed newborns, such as LCN2, CXCL1, CXCR2, S100A8 and S100A9, as well as histone-related proteins (e.g., HIST1H2BL). LCN2 is an iron-binding protein contained in neutrophil granules that is associated with neuroinflammation [69]. Using an animal model of chorioamnionitis, Yellowhair *et al.* [70] found that chorioamnionitis-induced fetal brain damage was associated with increased expression of CXCL1 and CXCR2 in placental and fetal brain, and with an elevated number of cerebral CXCR2<sup>+</sup> neutrophils. With respect to calgranulins, it has been observed that the brain expression of S100A8 and S100A9 proteins was increased in the brain of patients dying of sepsis; specifically, S100A9 expression was required for the cerebral recruitment of neutrophils and microglia activation in an animal model of sepsis [71].

Extracellular histones can also act as damage-associated molecular patterns, such as calgranulins, activating inflammatory signaling pathways and inducing multiple organ-specific damage [72]. In an animal study, histone H1 was found to induce pro-inflammatory responses in microglia and astrocytes [73,74]. We also found that both the gene and protein expression of S100A9, S100A8 and HIST1H2BL, were significantly higher in FIR-affected ELGANs, while our proteomics data also revealed the significant overexpression of HIST1H4A and HIST1H1C (Fig. 3C,D). In summary, our findings provide new insights into the molecular mechanisms that trigger ISSI after FIR and the onset of perinatal brain injury in ELGANs.

## Methods

### Study design

We performed a retrospective chart review of extreme preterm births (< 28 weeks of gestation) at the *Hospital Clínic de Barcelona* admitted to the neonatal intensive care unit between January 2010 and October 2016. Two hundred thirty-four cases born alive were found. From this initial

cohort, we excluded newborns from multiple pregnancies, outborn deliveries and those who presented major congenital anomalies, received a blood transfusion before collecting the DBS, had no histopathological examination of the placenta and umbilical cord, or whose DBS was collected later than 8 days after birth, most likely presenting molecular alterations unrelated to intrauterine conditions. To further avoid molecular alterations caused by severe organ dysfunction unrelated to FIR, we also excluded newborns who either died in the delivery room or within the first week after birth. We attempted to maximize the number of individuals with available biomarkers of intra-amniotic infection and inflammation in placenta and AF, while recruiting a similar number of FIR and non-FIR cases within a total maximum of 25, which as our results show, is sufficiently large to obtain a sizeable number of transcriptomic and proteomic changes. While this sample size may apparently be insufficient to enable a proper adjustment by relevant covariates, the high dimension of the molecular data allows one to circumvent that limitation by using surrogate variable analysis [75], as described later in this section. A final number of  $n = 21$  infants met the eligibility criteria and their families provided written consent to use their DBS samples for this study. At the time of processing these DBS samples, they had been stored at room temperature by the Catalan newborn screening program for between 1 and 7 years. The study protocol was approved by the Institutional Review Board from the *Hospital Clínic de Barcelona* (September 16<sup>th</sup>, 2016; ref. HCB/2016/0713).

### Clinical outcomes

GA at delivery was calculated according to the first-trimester ultrasound examination. Preterm birth was classified by the clinical presentation: threatened preterm labor, cervical insufficiency, preterm pre-labor rupture of membranes, placenta abruption, and preeclampsia. The laboratory results of AF cultures (genital mycoplasma, aerobic and anaerobic) and IL-6 levels, as well as the mother's blood CRP concentration, white blood cell (WBC) count and absolute neutrophil count (ANC) at admission were also recorded; see Table 2. Clinical chorioamnionitis was defined following the criteria of Gibbs, which includes presence of fever and any two of the following: maternal leukocytosis, maternal tachycardia, fetal tachycardia, uterine tenderness and foul-smelling vaginal discharge [76]. Slides from the placenta and umbilical cord were independently reviewed by two experienced pathologists blinded to experimental results and previous diagnosis. Maternal (MIR) and fetal (FIR) reactions were staged and graded according to published criteria [1,77]. Cases read discrepantly by the two pathologists were simultaneously re-reviewed on a two-headed microscope in order to reach a consensus (Table 1 and Fig. 1). The annotated neonatal morbidity included: respiratory distress syndrome (RDS); intraventricular



hemorrhage (IVH); white matter disease (WMD); retinopathy of prematurity (ROP); sepsis; necrotizing enterocolitis (NEC); patent ductus arteriosus (PDA); and laboratory investigations at birth, and during the first 8 days after birth (CRP concentration, WBC and ANC). Hypothesis tests on clinical variables used in Table 2 include two-tailed Fisher's exact test for categorical variables and *t*-tests for those that were continuous.

### RNA extraction, library preparation and sequencing

We used a scalpel to cut a surface of approximately 25 square millimeters from DBS samples in Guthrie cards, using a different scalpel blade for each sample to avoid contamination. RNA was then extracted using the illustra™ RNAspin Mini Isolation Kit (GE Healthcare, Chicago, IL, USA) following the supplier's protocol and eluting the sample in a final volume of 30  $\mu$ L. The quality and quantity of the RNA was assessed with an RNA Pico chip in an Agilent 2100 Bioanalyzer equipment. Next, globin mRNA and ribosomal RNA were removed using the Globin Zero Gold rRNA Removal Kit (Illumina), and its performance was checked with an RNA Pico chip in an Agilent 2100 Bioanalyzer equipment. Libraries were prepared using the NebNext Ultra™ II Directional RNA Library Prep Kit (New England Biolabs), following the specific protocol for rRNA depleted FFPE RNA and using 12 PCR cycles for library amplification. Libraries were validated and their concentration was measured with a High Sensitivity DNA chip in an Agilent 2100 Bioanalyzer. Finally, we prepared a pooled library with a normalized concentration for each sample. The final concentration of the pool was measured by Real-Time PCR using the NGS Library Quantification Kit (Takara, Mountain View, CA, USA). We loaded a final concentration of 1.9 pM into a NextSeq-500 High Output run with  $2 \times 75$  cycles using 4 sequencing lanes per library. The resulting  $21 \times 4 = 84$  paired-end FASTQ files have been deposited to the European Genome-Phenome Archive with the study identifier EGAS00001003635.

### Preprocessing and differential expression analysis of transcriptomics RNA-seq data

We performed quality control on the 84 raw paired-end reads using FastQC and did not detect artifacts in the sequence data. We did observe, however, that one of the samples, BS13, was sequenced at about 4 times more depth than the rest of the samples. The reason for this higher depth was a wrong normalization step during library preparation that led to a higher concentration of RNA in that sequenced library. To adjust for this bias in sequencing depth, we downsampled uniformly at random the paired-end reads from the library to  $\frac{1}{4}$  of its original depth.

Raw paired-end reads in FASTQ files, including the downsampled version of BS13, were aligned to the GRCh38 version of the reference human genome, without alternate locus scaffolds (GCA\_000001405.15) and including human decoy sequences from hs38d1 (GCA\_000786075.2), using STAR [78] version 2.6.0c with default parameters except for `--peOverlapNbasesMin 10` and `--sjdbOverhang 74`, producing an initial set of 84 BAM files. Aligned reads in BAM files were reduced to a table of counts of 25 221 genes by 84 samples using gene annotations from GENCODE v24 and the R/BIOCONDUCTOR package *GenomeAlignments* [79] version 1.16.0, and its function `summarizeOverlaps`. We used specific arguments in the call to this function to restrict the count of genic reads to only those that fell entirely within the exonic regions and aligned to a unique site on the genome, to reflect library preparation protocols and to avoid counting reads without a matching pair or overlapping multiple features.

The rest of the analysis of transcriptomics data was based on the *edgeR* [80] and *limma* [81] pipelines, and a few other R packages from the Bioconductor project [82] release version 3.7. A multidimensional scaling (MDS) plot of the samples showed a nearly perfect agreement between libraries from the same sample sequenced in each of the four different lanes. For this reason, we merged BAM files with aligned reads from the same sample and built a new table of counts of 25 221 genes by 21 samples, using the previously employed reduction procedure. Considering that 6 out of the 21 infants were not diagnosed with MIR or FIR (Fig. 1), and following previously established recommendations [83], we filtered out lowly expressed genes by discarding those that did not show a minimum reliable level of expression of 10 counts per million reads of the smallest library size, in at least 6 samples. This filtering step led to a final table of counts of 11 279 genes by 21 samples.

We used TMM normalization [84] to produce the MDS plot shown in Fig. 2C, as well as to obtain gene-level robust dispersion estimates [85] that helped to select the sex of the infant and surrogate variables calculated with SVA [75], as covariates in the linear models employed for differential expression analysis. Finally, we used the *limma-voom* [86] pipeline with quality weights, combined with quantile normalization, to conduct the differential expression analysis using linear models. The design matrix included the FIR status of the infants as a main explanatory variable and sex and surrogate variables, estimated with SVA, as covariates.

The differential expression analysis across different combined stages of inflammation in MIR and FIR (Fig. 9, Table S3) was based on our replacing the FIR status variable in the linear model, with numerical values for MIR and FIR. Specifically, we assigned a 0 when MIR and FIR were not present, 1 when MIR was present but FIR was not, 2 when MIR was stage 3 and FIR was stage 1 and, finally, 3 when MIR was stage 3 and FIR was stage 2 or 3.

## Preprocessing and differential expression analysis of transcriptomics microarray data

We obtained from Nigel Paneth and collaborators at the Michigan State University raw microarray gene expression data files from paired frozen and unfrozen DBS samples from nine neonates of the US ELGAN cohort, where two were affected by FIR and seven were not. The  $n = 18$  corresponding samples had been hybridized on a Agilent SurePrint G3 Human Gene Expression 8x60K Microarray Kit, and their analysis comparing the number of expressed genes between frozen and unfrozen samples was published by Wei *et al.* [15]. Batch processing information was derived from the scanning timestamp stored in raw data files as previously described [87], creating a batch indicator variable that divided the samples into two different batches. The cross-classification of infants by FIR status and batch indicator showed no correlation between the primary outcome of our analysis, FIR status, and sample batch processing (Fig. 4 and 6). The MDS plot of the data shows the effect of storage temperature and batch processing, but not sex (Fig. 6).

We corrected background expression intensities using the 'normexp' method in the function backgroundCorrect from the *limma* package [88], and used a quantile normalization approach to obtain comparable gene expression distributions across the samples (Fig. 5). We removed probesets without annotation to Entrez gene identifiers, and probesets annotated to a common Entrez gene, keeping the one with highest variability measured by interquartile range (IQR). We then discarded the Agilent positive control probes and retained initially the negative control probes to later estimate the fraction of expressed genes per sample. This resulted in an expression data set of 21 952 probesets in one-to-one correspondence with Entrez genes after discarding negative control probes, by  $n = 18$  samples. We calculated the 95% quantile of the expression level distribution observed in negative control probes for each sample. As a function of the percentage of increase over the 95% quantile of negative probe expression (from 100% to 200% increasing by 10%), we found the number of expressed genes per sample separately by storage condition, which we later used for nonspecific filtering based on minimum expression.

Using the software package *sva* [75] and the top 1% most variable (IQR) genes, we estimated one surrogate variable covariate that captures variability unrelated to FIR, storage temperature and batch processing. For each gene, we defined a linear mixed-effects model where the expression profile of every gene is a linear function of FIR status (yes/no, main explanatory variable), storage temperature (frozen/unfrozen), batch processing, the estimated surrogate variable covariate and the infant as a random effect. We estimated array quality weights to down-weight samples of worse quality in a graduated way [89] and to avoid

discarding them. Using *limma*, we fitted this linear mixed-effects model, including the array quality weights, to each gene expression profile. We then calculated moderated  $t$ -statistics, squeezing the gene-wise residual variances toward a global trend, for the coefficient estimating the effect between FIR and non-FIR-affected infants, and their corresponding  $P$  values for the null hypothesis of no-differential expression (Fig. 7). We selected initially different subsets of DE genes at 5% FDR according to different thresholds of the nonspecific filtering on minimum expression and, within each nonspecific filter, FDR-adjusted  $P$  values were re-calculated. We found that a 110% cutoff yielded the highest number of DE genes and using this nonspecific filter we readjusted the raw  $P$  values and obtained 33 genes that changed significantly at 5% FDR by 50% or more, between newborns affected and unaffected by FIR (Figs 3B and 7B and Table S2).

## Mass spectrometry proteomics

### Sample preparation for mass spectrometry

DBS samples were excised and collected in a 2 mL Eppendorf, and extracted as described previously [90]. Briefly, each sample was soaked with 970  $\mu$ L of a solution consisting of 758  $\mu$ L of 25 mM ammonium bicarbonate, 113  $\mu$ L of 10% (w/v) sodium deoxycholate and 99  $\mu$ L of 5 mM tris(2-carboxyethyl)phosphine, which was incubated at 60 °C during 1 h in constant agitation. Samples were alkylated with 10  $\mu$ M iodoacetamide (37 °C, 30 min) and quenched with 10  $\mu$ M dithiothreitol (37 °C, 30 min). Samples were digested with trypsin (40 ng- $\mu$ L<sup>-1</sup>) overnight at 37 °C. Digestion was stopped with formic acid, and samples were vortexed and centrifuged (800  $\times$  g 20 min) prior to supernatant desalting using spin C18 columns.

### Mass spectrometry data acquisition

Samples were reconstituted in 0.1% formic acid and were analyzed on a Orbitrap Fusion Lumos with an EASY-Spray nanosource coupled to a nano-UPLC system (EASY-nanoLC 1000 liquid chromatograph, Thermo-Fisher, San Jose, CA, USA) equipped with a 50-cm C18 column (EASY-Spray; 75  $\mu$ m id, PepMap RSLC C18, 2 mm particles, 45 °C). Chromatographic gradients were started at 5% buffer B with a flow rate of 300 nL-min<sup>-1</sup> and gradually increased to 22% buffer B in 79 min and to 32% in 11 min. After each analysis, the column was washed for 10 min with 95% buffer B (Buffer A: 0.1% formic acid in water. Buffer B: 0.1% formic acid in acetonitrile). The mass spectrometer was operated in data-dependent acquisition mode, with full MS scans over a mass range of  $m/z$  350–1500 with detection in the Orbitrap (120 K resolution) and with auto gain control (AGC) set to 100 000. In each cycle of data-dependent acquisition

analysis, following each survey scan, the most intense ions above a threshold ion count of 10 000 were selected for fragmentation with HCD at normalized collision energy of 28%. The number of selected precursor ions for fragmentation was determined by the 'Top Speed' acquisition algorithm (max cycle time of 3 s), and a dynamic exclusion of 60 s was set. Fragment ion spectra were acquired in the ion trap with an AGC of 10 000 and a maximum injection time of 200 ms. One of the samples, BS04, could not be profiled for technical reasons. All data were acquired with XCALIBUR software v4.1 (ThermoFisher).

### Mass spectrometry protein expression quantification

The MAXQUANT software suite (v1.6.2.6) was used for peptide identification and label-free protein quantification [91]. The data were searched against the UniProt human database with decoy entries (as of April 2018, 42 518 entries). A precursor ion mass tolerance of 4.5 ppm at the MS1 level was used, and up to two missed cleavages for trypsin were allowed. The fragment ion mass tolerance was set to 0.5 Da. Oxidation of methionine and protein acetylation at the N-terminal were defined as variable modification, whereas carbamidomethylation on cysteines was set as a fixed modification. The minimum number of razor and unique peptides ('min. razor + unique') for a protein group to be considered as identified was set to 1. Identified peptides and proteins were filtered using a 5% FDR. The match between runs algorithm was activated with a tolerance of 0.7 min for the matching time window and 20 min for the alignment time window. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [92] partner repository with accession number PXD011626.

### Preprocessing and differential expression analysis of proteomics data

From the initial set of 649 quantified protein expression profiles on 20 samples, we discarded 26 whose protein identifier could not be mapped to one of the 25 221 genes profiled by RNA-seq. From the remaining 624 proteins, we selected 303 with protein quantification values in at least 6 out of the 20 samples. We further discarded 48 proteins whose protein quantification values were based on one single peptide, leading to a final protein quantification data matrix of 245 proteins by 20 samples. Raw protein quantifications were normalized using the *vsN* [93,94] package. A complete normalized matrix of protein expression values was obtained by imputing missing quantifications with a multivariate method [95] based on the expectation-maximization algorithm implemented in the function *impute.wrapper.MLE* from the R package *imputeLCMD*.

Using this complete and normalized matrix of protein expression values as an input to the *edgeR* [80] and *limma-trend* [85] pipelines, we produced the MDS plot in Fig. 2D and conducted a protein differential expression analysis with linear models shown in Fig. 3C. The design matrix included the FIR status of infants as the main explanatory variable and sex as a covariate.

### Functional enrichment analysis with the Gene Ontology database

We performed functional enrichment analysis with the GO database using the package *GOstats* [96]. More concretely, we used the biological process ontology and the conditional hypergeometric test to account for dependencies derived from the GO hierarchical structure. Under this procedure, the GO hierarchy is traversed toward the root, applying a one-tailed Fisher's exact test between the genes in the GO term and the DE gene set at hand. When a GO term is considered to be significantly enriched by the DE gene set, its genes are removed from parent GO terms before testing them. To consider a GO term as significantly enriched we used a *P* value cutoff of 0.01, a minimum OR of 1.5, a minimum of five and a maximum of 300 genes in the GO term. The gene universe was defined by the entire set of 25 221 profiled genes by RNA-seq. To report only the most robustly enriched biological processes, the final set of significantly enriched GO terms was selected by first using a multiple testing correction of FDR < 10% on the whole set of tested GO terms, then by further selecting GO terms with OR > 1.5 and a minimum number of five genes enriching the GO term.

### Analysis of T-cell receptor excision circles

TREC data were obtained using the EnLite™ Neonatal TREC kit (PerkinElmer, Turku, Finland), which is a combination of PCR-based nucleic acid amplification and time-resolved fluorescence resonance energy transfer based detection. EnLite™ simultaneously detects two targets, TREC and beta-actin, the latter used as an internal control for monitoring specimen amplification in each test. The assay involves DNA elution from 1.5 mm DBS punches, amplification and hybridization with target-sequence specific probes, and quantifies TREC levels by measuring probe fluorescence with the Victor Enlite™ fluorometer (PerkinElmer). A full calibration curve with blanks and three DBS calibrators was run in triplicate on each plate. A low-TREC control, a no-TREC control, a high-TREC control, and a blank paper disk (containing no blood), were used as quality controls for each plate. Raw fluorescence counts, measured at 615, 665, and 780 nm, were processed by the EnLite™ workstation software to produce corrected fluorescence counts. TREC and

beta-actin copies per  $\mu\text{L}$  were predicted from corrected fluorescent counts by the EnLite™ workstation software using an unweighted linear regression model on the ArcSinh transformation of the DBS calibrator copies per  $\mu\text{L}$  response.

The values for the number of TREC copies per  $\mu\text{L}$  are sensitive to technical effects associated with the run in which they are obtained; for this reason, we generated technical replicates for the same individuals in different runs. Because of limited DBS availability, we could not produce more than one TREC readout for all individuals and not all individuals with multiple readouts had the same number of replicated measurements. For this reason, we used a linear mixed-effects model to test the effect of FIR on the TREC readout, using the newborn identifier as a random effect variable to account for replicated measurements across individuals. We used the logarithm of the corrected fluorescent counts as TREC readout instead of the predicted number of copies per  $\mu\text{L}$ , as these more accurately reflect the measurements made by the instrument and its technical biases. In this linear mixed-effects model, the FIR stage was the main explanatory variable, and aside from the random effect of the newborn, other covariates that entered as main factors in the model were sex, maximum lymphocyte count during the first postnatal week, and a batch indicator variable identifying the run of the TREC measurement. This linear mixed-effects model, as well as a null one without the FIR stage, was fitted to the data using the package *lme4* [97].

A chi-squared test with one degree of freedom between the two models gave a  $P$  value = 0.0432, thus rejecting the hypothesis of no FIR effect at a significance level  $\alpha < 0.05$ . Adjusted and corrected fluorescent counts on a logarithmic scale (shown in Fig. 10A) were obtained after removing the effect of the previous covariates using a linear model.

## Acknowledgements

This work was supported by the Spanish MINECO/FEDER (TIN2015-71079-P). The CRG/UPF Proteomics Unit is part of the Spanish Infrastructure for Omics Technologies (ICTS Omics Tech) and is a member of Proteored, PRB3, which is supported by grant PT17/0019 of the PE I + D+i 2013-2016, funded by ISCIII and ERDF. We acknowledge support from the Spanish Ministry of Economy and Competitiveness, ‘Centro de Excelencia Severo Ochoa 2013-2017’ (SEV-2012-0208), ‘Unidad de Excelencia María de Maeztu’ (CEX2018-000792-M), and ‘Secretaria d’Universitats i Recerca del Departament d’Economia i Coneixement de la Generalitat de Catalunya’ (SGR17-595; SGR17-1020), and the Spanish National Bioinformatics Institute-ISCIII (PT17/0009/0014). The authors thank the ELGAN PAD committee, and Nigel Paneth, Sok

Kean Khoo, Pete Haak, Madeleine Lenski, and Elizabeth Allred, for granting access to the microarray expression and clinical data from the US ELGAN cohort. We would like to thank Aida Andrés, José Aramburu, Sergi Castellano, Alan Leviton, and Cristina López-Rodríguez for their critical remarks on various parts of the manuscript.

## Conflict of interest

The authors declare no conflict of interest.

## Author contributions

DC, TC, and RC conceived the study. DC, NB, AS, JMG, ES, FC, CR, AN, and JLM collected the data. DC and RC analyzed the data. DC, TC, and RC drafted the manuscript. All authors revised and approved the manuscript.

## Ethics approval

The study was approved by the Institutional Review Board from the Hospital Clínic de Barcelona (September 16th, 2016; ref. HCB/2016/0713). The parents of the analyzed infants provided written, informed consent.

## Data accessibility

All raw data is deposited at the EGA and PRIDE databases. The processed RNA-seq data is deposited along with the raw RNA-seq data at the EGA database. The processed raw protein expression profiles are included as Supplementary Data File S1.

The transcriptomics and clinical data reported in this paper are available through the European Genome-phenome Archive (EGA) under accession number EGAS00001003635. The mass spectrometry proteomics data are available through the PRIDE repository under accession number PXD011626.

## References

- 1 Khong TY, Mooney EE, Ariel I, Balmus NCM, Boyd TK, Brundler M-A, Derricott H, Evans MJ, Faye-Petersen OM, Gillan JE *et al.* (2016) Sampling and definitions of placental lesions: Amsterdam Placental Workshop Group consensus statement. *Arch Pathol Lab Med* **140**, 698–713.
- 2 Gantert M, Been JV, Gavilanes AWD, Garnier Y, Zimmermann LJI & Kramer BW (2010) Chorioamnionitis: a multiorgan disease of the fetus? *J Perinatol* **30** (Suppl), S21–S30.

- 3 Dammann O & Leviton A (2000) Role of the fetus in perinatal infection and neonatal brain damage. *Curr Opin Pediatr* **12**, 99–104.
- 4 Costa D & Castelo R (2016) Umbilical cord gene expression reveals the molecular architecture of the fetal inflammatory response in extremely preterm newborns. *Pediatr Res* **79**, 473–481.
- 5 Tilley SK, Joseph RM, Kuban KCK, Dammann OU, O'Shea TM & Fry RC (2017) Genomic biomarkers of prenatal intrauterine inflammation in umbilical cord tissue predict later life neurological outcomes. *PLoS One* **12**, e0176953.
- 6 Leviton A, Hecht JL, Allred EN, Yamamoto H, Fichorova RN, Dammann O & ELGAN Study Investigators (2011) Persistence after birth of systemic inflammation associated with umbilical cord inflammation. *J Reprod Immunol* **90**, 235–243.
- 7 Jackson CM, Wells CB, Tabangin ME, Meinzen-Derr J, Jobe AH & Choungnet CA (2017) Pro-inflammatory immune responses in leukocytes of premature infants exposed to maternal chorioamnionitis or funisitis. *Pediatr Res* **81**, 384–390.
- 8 O'Shea TM, Shah B, Allred EN, Fichorova RN, Kuban KCK, Dammann O, Leviton A & ELGAN Study Investigators (2013) Inflammation-initiating illnesses, inflammation-related proteins, and cognitive impairment in extremely preterm infants. *Brain Behav Immun* **29**, 104–112.
- 9 Dammann O & Leviton A (2014) Intermittent or sustained systemic inflammation and the preterm brain. *Pediatr Res* **75**, 376–380.
- 10 Dammann O, Allred EN, Fichorova RN, Kuban K, O'Shea TM, Leviton A & ELGAN Study Investigators (2016) Duration of systemic inflammation in the first postnatal month among infants born before the 28th week of gestation. *Inflammation* **39**, 672–677.
- 11 Yanni D, Korzeniewski SJ, Allred EN, Fichorova RN, O'Shea TM, Kuban K, Dammann O & Leviton A (2017) Both antenatal and postnatal inflammation contribute information about the risk of brain damage in extremely preterm newborns. *Pediatr Res* **82**, 691–696.
- 12 Korzeniewski SJ, Romero R, Cortez J, Pappas A, Schwartz AG, Kim CJ, Kim J-S, Kim YM, Yoon BH, Chaiworapongsa T *et al.* (2014) A “multi-hit” model of neonatal white matter injury: cumulative contributions of chronic placental inflammation, acute fetal inflammation and postnatal inflammatory events. *J Perinat Med* **42**, 731–743.
- 13 Papile LA, Burstein J, Burstein R & Koffler H (1978) Incidence and evolution of subependymal and intraventricular hemorrhage: a study of infants with birth weights less than 1,500 gm. *J Pediatr* **92**, 529–534.
- 14 Sawyer S, Krause J, Guschanski K, Savolainen V & Pääbo S (2012) Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One* **7**, e34131.
- 15 Wei C, Lu Q, Khoo SK, Lenski M, Fichorova RN, Leviton A & Paneth N (2014) Comparison of frozen and unfrozen blood spots for gene expression studies. *J Pediatr* **164**, 189–191.e1.
- 16 Canna SW, Girard C, Malle L, de Jesus A, Romberg N, Kelsen J, Surrey LF, Russo P, Sleight A, Schiffrin E *et al.* (2017) Life-threatening NLRC4-associated hyperinflammation successfully treated with IL-18 inhibition. *J Allergy Clin Immunol* **139**, 1698–1701.
- 17 Condamine T, Dominguez GA, Youn J-I, Kossenkov AV, Mony S, Alicea-Torres K, Tcyganov E, Hashimoto A, Nefedova Y, Lin C *et al.* (2016) Lectin-type oxidized LDL receptor-1 distinguishes population of human polymorphonuclear myeloid-derived suppressor cells in cancer patients. *Sci Immunol* **1**, aaf8943.
- 18 Breuer K, Foroushani AK, Laird MR, Chen C, Sribnaia A, Lo R, Winsor GL, Hancock REW, Brinkman FSL & Lynn DJ (2013) InnateDB: systems biology of innate immunity and beyond—recent updates and continuing curation. *Nucleic Acids Res* **41**, D1228–D1233.
- 19 Naranbhai V, Fairfax BP, Makino S, Humburg P, Wong D, Ng E, Hill AVS & Knight JC (2015) Genomic modulators of gene expression in human neutrophils. *Nat Commun* **6**, 7545.
- 20 Boyman O, Krieg C, Homann D & Sprent J (2012) Homeostatic maintenance of T cells and natural killer cells. *Cell Mol Life Sci* **69**, 1597–1608.
- 21 Haines CJ, Giffon TD, Lu L-S, Lu X, Tessier-Lavigne M, Ross DT & Lewis DB (2009) Human CD4 T cell recent thymic emigrants are identified by protein tyrosine kinase 7 and have reduced immune function. *J Exp Med* **206**, 275–285.
- 22 Lewis DB, Haines C & Ross D (2011) Protein tyrosine kinase 7: a novel surface marker for human recent thymic emigrants with potential clinical utility. *J Perinatol* **31** (Suppl 1), S72–S81.
- 23 Pérès M, Montfort A, Andrieu-Abadie N, Colacios C & Ségui B (2017) S1P: the elixir of life for naive T cells. *Cell Mol Immunol* **15**, 657–659.
- 24 Glavina-Durdov M, Springer O, Capkun V, Saratlija-Novaković Z, Rozić D & Barle M (2003) The grade of acute thymus involution in neonates correlates with the duration of acute illness and with the percentage of lymphocytes in peripheral blood smear. Pathological study. *Biol Neonate* **83**, 229–234.
- 25 Schwarz J, Scheckenbach V, Kugel H, Spring B, Pagel J, Härtel C, Pauluschke-Fröhlich J, Peter A, Poets CF, Gille C *et al.* (2018) Granulocytic myeloid-derived suppressor cells (GR-MDSC) accumulate in cord blood of preterm infants and remain elevated during the neonatal period. *Clin Exp Immunol* **191**, 328–337.



- 26 Gervassi A, Lejarcegui N, Dross S, Jacobson A, Itaya G, Kidzeru E, Gantt S, Jaspan H & Horton H (2014) Myeloid derived suppressor cells are present at high frequency in neonates and suppress in vitro T cell responses. *PLoS One* **9**, e107816.
- 27 Leiber A, Schwarz J, Köstlin N, Spring B, Fehrenbach B, Katava N, Poets CF & Gille C (2017) Neonatal myeloid derived suppressor cells show reduced apoptosis and immunosuppressive activity upon infection with *Escherichia coli*. *Eur J Immunol* **47**, 1009–1021.
- 28 Schieber M & Chandel NS (2014) ROS function in redox signaling and oxidative stress. *Curr Biol* **24**, R453–R462.
- 29 Veglia F, Perego M & Gabrilovich D (2018) Myeloid-derived suppressor cells coming of age. *Nat Immunol* **19**, 108–119.
- 30 Zhou J, Nefedova Y, Lei A & Gabrilovich D (2018) Neutrophils and PMN-MDSC: their biological role and interaction with stromal cells. *Semin Immunol* **35**, 19–28.
- 31 Ost M, Singh A, Peschel A, Mehling R, Rieber N & Hartl D (2016) Myeloid-Derived Suppressor Cells in Bacterial Infections. *Front Cell Infect Microbiol* **6**, 37
- 32 Tamadaho RSE, Hoerauf A & Layland LE (2018) Immunomodulatory effects of myeloid-derived suppressor cells in diseases: role in cancer and infections. *Immunobiology* **223**, 432–442.
- 33 Moskovitz J (2005) Roles of methionine sulfoxide reductases in antioxidant defense, protein regulation and survival. *Curr Pharm Des* **11**, 1451–1457.
- 34 Duncan JA & Canna SW (2017) The NLR4 inflammasome. *Immunol Rev* **281**, 115–123.
- 35 Wynn JL, Wilson CS, Hawiger J, Scumpia PO, Marshall AF, Liu J-H, Zharkikh I, Wong HR, Lahni P, Benjamin JT *et al.* (2016) Targeting IL-17A attenuates neonatal sepsis mortality induced by IL-18. *Proc Natl Acad Sci USA* **113**, E2627–E2635.
- 36 Liang J, Alfano DN, Squires JE, Riley MM, Parks WT, Kofler J, El-Gharbawy A, Madan-Kheterpal S, Acquaro R & Picarsic J (2017) Novel NLR4 mutation causes a syndrome of perinatal autoinflammation with hemophagocytic lymphohistiocytosis, hepatosplenomegaly, fetal thrombotic vasculopathy, and congenital anemia and ascites. *Pediatr Dev Pathol* **20**, 498–505.
- 37 Denes A, Coutts G, Lénárt N, Cruickshank SM, Pelegrin P, Skinner J, Rothwell N, Allan SM & Brough D (2015) AIM2 and NLR4 inflammasomes contribute with ASC to acute brain injury independently of NLRP3. *Proc Natl Acad Sci USA* **112**, 4050–4055.
- 38 Ravichandran S, Chouthai NS, Patel B, Sharma A, Gupte A, Ma MM, Mamilla D, Lulic-Botica M, Thomas R & Kamat D (2019) Higher daily doses of caffeine lowered the incidence of moderate to severe neurodevelopmental disabilities in very low birth weight infants. *Acta Paediatr* **108**, 430–435.
- 39 Furman D, Chang J, Lartigue L, Bolen CR, Haddad F, Gaudilliere B, Ganio EA, Fragiadakis GK, Spitzer MH, Douchet I *et al.* (2017) Expression of specific inflammasome gene modules stratifies older individuals into two extreme clinical and immunological states. *Nat Med* **23**, 174–184.
- 40 Schett G & Neurath MF (2018) Resolution of chronic inflammatory disease: universal and tissue-specific concepts. *Nat Commun* **9**, 3261.
- 41 Hanna N, Vasquez P, Pham P, Heck DE, Laskin JD, Laskin DL & Weinberger B (2005) Mechanisms underlying reduced apoptosis in neonatal neutrophils. *Pediatr Res* **57**, 56–62.
- 42 Allgaier B, Shi M, Luo D & Koenig JM (1998) Spontaneous and Fas-mediated apoptosis are diminished in umbilical cord blood neutrophils compared with adult neutrophils. *J Leukoc Biol* **64**, 331–336.
- 43 Nguyen CN, Schnulle PM, Chegini N, Luo X & Koenig JM (2010) Neonatal neutrophils with prolonged survival secrete mediators associated with chronic inflammation. *Neonatology* **98**, 341–347.
- 44 Atallah M, Krispin A, Trahtemberg U, Ben-Hamron S, Grau A, Verbovetski I & Mevorach D (2012) Constitutive neutrophil apoptosis: regulation by cell concentration via S100 A8/9 and the MEK – ERK pathway. *PLoS One* **7**, e29333.
- 45 Hirata J-I, Kotani J, Aoyama M, Kashiwamura S-I, Ueda H, Kuroda Y, Usami M, Okamura H & Marukawa S (2008) A role for IL-18 in human neutrophil apoptosis. *Shock* **30**, 628–633.
- 46 Chen KW, Groß CJ, Sotomayor FV, Stacey KJ, Tschopp J, Sweet MJ & Schroder K (2014) The neutrophil NLR4 inflammasome selectively promotes IL-1 $\beta$  maturation without pyroptosis during acute Salmonella challenge. *Cell Rep* **8**, 570–582.
- 47 Kessel C, Holzinger D & Foell D (2013) Phagocyte-derived S100 proteins in autoinflammation: putative role in pathogenesis and usefulness as biomarkers. *Clin Immunol* **147**, 229–241.
- 48 Alder MN, Opoka AM, Lahni P, Hildeman DA & Wong HR (2017) Olfactomedin-4 is a candidate marker for a pathogenic neutrophil subset in septic shock. *Crit Care Med* **45**, e426–e432.
- 49 Solan PD, Dunsmore KE, Denenberg AG, Odoms K, Zingarelli B & Wong HR (2012) A novel role for matrix metalloproteinase-8 in sepsis. *Crit Care Med* **40**, 379–387.
- 50 O'Hare FM, Watson RWG, O'Neill A, Blanco A, Donoghue V & Molloy EJ (2016) Persistent systemic monocyte and neutrophil activation in neonatal encephalopathy. *J Matern Fetal Neonatal Med* **29**, 582–589.

- 51 Molloy EJ, O'Neill AJ, Grantham-Sloan JJ, Sheridan-Pereira M, Fitzpatrick JM, Webb DW & Watson RWG (2007) Neonatal encephalopathy is associated with altered perinatal systemic neutrophil apoptosis. *Am J Perinatol* **24**, 525–530.
- 52 Venet F & Monneret G (2018) Advances in the understanding and treatment of sepsis-induced immunosuppression. *Nat Rev Nephrol* **14**, 121–137.
- 53 Hibbert JE, Currie A & Strunk T (2018) Sepsis-induced immunosuppression in neonates. *Front Pediatr* **6**, 357.
- 54 Toti P, De Felice C, Occhini R, Schuerfeld K, Stumpo M, Epistolato MC, Vatti R & Buonocore G (2004) Spleen depletion in neonatal sepsis and chorioamnionitis. *Am J Clin Pathol* **122**, 765–771.
- 55 Toti P, De Felice C, Stumpo M, Schürfeld K, Di Leo L, Vatti R, Bianciardi G, Buonocore G, Seemayer TA & Luzi P (2000) Acute thymic involution in fetuses and neonates with chorioamnionitis. *Hum Pathol* **31**, 1121–1128.
- 56 Azizia M, Lloyd J, Allen M, Klein N & Peebles D (2012) Immune status in very preterm neonates. *Pediatrics* **129**, e967–e974.
- 57 Kamdar S, Hutchinson R, Laing A, Stacey F, Ansbro K, Millar MR, Costeloe K, Wade WG, Fleming P & Gibbons DL (2020) Perinatal inflammation influences but does not arrest rapid immune development in preterm babies. *Nat Commun* **11**, 1284.
- 58 de Jong E, Hancock DG, Wells C, Richmond P, Simmer K, Burgner D, Strunk T & Currie AJ (2018) Exposure to chorioamnionitis alters the monocyte transcriptional response to the neonatal pathogen *Staphylococcus epidermidis*. *Immunol Cell Biol* **96**, 792–804.
- 59 Bermick J, Gallagher K, denDekker A, Kunkel S, Lukacs N & Schaller M (2019) Chorioamnionitis exposure remodels the unique histone modification landscape of neonatal monocytes and alters the expression of immune pathway genes. *FEBS J* **286**, 82–109.
- 60 Di Naro E, Cromi A, Ghezzi F, Raio L, Uccella S, D'Addario V & Loverro G (2006) Fetal thymic involution: a sonographic marker of the fetal inflammatory response syndrome. *Am J Obstet Gynecol* **194**, 153–159.
- 61 Scheible KM, Emo J, Laniewski N, Baran AM, Peterson DR, Holden-Wiltse J, Bandyopadhyay S, Straw AG, Huyck H, Ashton JM *et al.* (2018) T cell developmental arrest in former premature infants increases risk of respiratory morbidity later in infancy. *JCI Insight* **3**, e96724.
- 62 Kuban JD, Allred EN & Leviton A (2006) Thymus involution and cerebral white matter damage in extremely low gestational age neonates. *Biol Neonate* **90**, 252–257.
- 63 He Y-M, Li X, Perego M, Nefedova Y, Kossenkov AV, Jensen EA, Kagan V, Liu Y-F, Fu S-Y, Ye Q-J *et al.* (2018) Transitory presence of myeloid-derived suppressor cells in neonates is critical for control of inflammation. *Nat Med* **24**, 224–231.
- 64 Chemnitz JM, Driesen J, Classen S, Riley JL, Debey S, Beyer M, Popov A, Zander T & Schultze JL (2006) Prostaglandin E2 impairs CD4+ T cell activation by inhibition of I $\kappa$ B: implications in Hodgkin's lymphoma. *Cancer Res* **66**, 1114–1122.
- 65 Ozsürekci Y & Aykac K (2016) Oxidative stress related diseases in newborns. *Oxid Med Cell Longev* **2016**, 1–9.
- 66 Tindell R & Tipple T (2018) Selenium: implications for outcomes in extremely preterm infants. *J Perinatol* **38**, 197–202.
- 67 Fong ON, Chan KYY, Leung KT, Lam HS, Cheung HM, Leung TY, Li K & Ng PC (2014) Expression profile of cord blood neutrophils and dysregulation of HSPA1A and OLR1 upon challenge by bacterial peptidoglycan. *J Leukoc Biol* **95**, 169–178.
- 68 Kruzel ML, Zimecki M & Actor JK (2017) Lactoferrin in a context of inflammation-induced pathology. *Front Immunol* **8**.
- 69 Song J & Kim OY (2018) Perspectives in lipocalin-2: emerging biomarker for medical diagnosis and prognosis for Alzheimer's disease. *Clin Nutr Res* **7**, 1–10.
- 70 Yellowhair TR, Noor S, Maxwell JR, Anstine CV, Oppong AY, Robinson S, Milligan ED & Jantzie LL (2018) Preclinical chorioamnionitis dysregulates CXCL1/CXCR2 signaling throughout the placental-fetal-brain axis. *Exp Neurol* **301**, 110–119.
- 71 Denstaedt SJ, Spencer-Segal JL, Newstead MW, Laborc K, Zhao AP, Hjelmaas A, Zeng X, Akil H, Standiford TJ & Singer BH (2018) S100A8/A9 drives neuroinflammatory priming and protects against anxiety-like behavior after sepsis. *J Immunol* **200**, 3188–3200.
- 72 Silk E, Zhao H, Weng H & Ma D (2017) The role of extracellular histone in organ injury. *Cell Death Dis* **8**, e2812.
- 73 Gilthorpe JD, Oozeer F, Nash J, Calvo M, Bennett DL, Lumsden A & Pini A (2013) Extracellular histone H1 is neurotoxic and drives a pro-inflammatory response in microglia. *F1000Res* **2**, 148.
- 74 Hagberg H, Mallard C, Ferriero DM, Vannucci SJ, Levison SW, Vexler ZS & Gressens P (2015) The role of inflammation in perinatal brain injury. *Nat Rev Neurol* **11**, 192–208.
- 75 Leek JT & Storey JD (2007) Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet* **3**, 1724–1735.
- 76 Gibbs RS & Duff P (1991) Progress in pathogenesis and management of clinical intraamniotic infection. *Am J Obstet Gynecol* **164**, 1317–1326.
- 77 Redline RW, Faye-Petersen O, Heller D, Qureshi F, Savell V, Vogler C & Society for Pediatric Pathology, Perinatal Section, Amniotic Fluid Infection Nosology Committee (2003) Amniotic infection syndrome:

- nosology and reproducibility of placental reaction patterns. *Pediatr Dev Pathol* **6**, 435–448.
- 78 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M & Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21.
- 79 Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, Morgan MT & Carey VJ (2013) Software for computing and annotating genomic ranges. *PLoS Comput Biol* **9**, e1003118.
- 80 Robinson MD, McCarthy DJ & Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140.
- 81 Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W & Smyth GK (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47.
- 82 Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, Bravo HC, Davis S, Gatto L, Girke T *et al.* (2015) Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods* **12**, 115–121.
- 83 Chen Y, Lun ATL & Smyth GK (2016) From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Res* **5**, 1438.
- 84 Robinson MD & Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**, R25.
- 85 Phipson B, Lee S, Majewski IJ, Alexander WS & Smyth GK (2016) Robust hyperparameter estimation protects against hypervariable genes and improves power to detect differential expression. *Ann Appl Stat* **10**, 946–963.
- 86 Law CW, Chen Y, Shi W & Smyth GK (2014) voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* **15**, R29.
- 87 Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, Johnson WE, Geman D, Baggerly K & Irizarry RA (2010) Tackling the widespread and critical impact of batch effects in high-throughput data. *Nat Rev Genet* **11**, 733–739.
- 88 Silver JD, Ritchie ME & Smyth GK (2009) Microarray background correction: maximum likelihood estimation for the normal-exponential convolution. *Biostatistics* **10**, 352–363.
- 89 Ritchie ME, Diyagama D, Neilson J, van Laar R, Dobrovic A, Holloway A & Smyth GK (2006) Empirical array quality weights in the analysis of microarray data. *BMC Bioinformatics* **7**, 261.
- 90 Chambers AG, Percy AJ, Yang J & Borchers CH (2015) Multiple reaction monitoring enables precise quantification of 97 proteins in dried blood spots. *Mol Cell Proteomics* **14**, 3094–3104.
- 91 Cox J & Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367–1372.
- 92 Vizcaíno JA, Côté RG, Csordas A, Dianes JA, Fabregat A, Foster JM, Griss J, Alpi E, Birim M, Contell J *et al.* (2013) The PRoteomics IDentifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res* **41**, D1063–D1069.
- 93 Huber W, von Heydebreck A, Sültmann H, Poustka A & Vingron M (2002) Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* **18**(Suppl 1), S96–S104.
- 94 Välikangas T, Suomi T & Elo LL (2018) A systematic evaluation of normalization methods in quantitative label-free proteomics. *Brief Bioinform* **19**, 1–11.
- 95 Schafer JL (1997) Analysis of Incomplete Multivariate Data. Chapman & Hall / CRC, New York. ISBN 0-412-04061-1.
- 96 Falcon S & Gentleman R (2007) Using GOstats to test gene lists for GO term association. *Bioinformatics* **23**, 257–258.
- 97 Bates D, Mächler M, Bolker B & Walker S (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* **67**, 1–48.

## Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Table S1.** Differential expression analysis of the FIR effect.

**Table S2.** Differential expression analysis of the FIR effect in the US ELGAN cohort.

**Table S3.** Differential expression analysis of the combined linear effect of MIR and FIR.

**Table S4.** Differential protein expression analysis of the FIR effect.

**Table S5.** Gene ontology (GO) analysis of FIR upregulated genes.

**Table S6.** Differentially expressed genes in FIR involved in neutrophil biology (Borregaard *et al.*, 2007; Futosi *et al.*, 2013; Naranbai *et al.*, 2015).

**Table S7.** Downregulated genes in FIR (FDR < 5%, minimum 1.5-fold change larger in nonFIR neonates) that code for transcription factors and significantly enrich this functional category.

**Table S8.** Gene ontology (GO) analysis of FIR downregulated genes.

**Table S9.** Comparison of FIR gene expression signatures from umbilical cord tissue and dried blood spots.

**Data S1.** Raw protein expression profiles.