

## 2 **Molecular profiling of neonatal dried blood spots reveals changes in** 3 **innate and adaptive immunity following fetal inflammatory response**

4  
5 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>,  
6 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>6</sup>, Teresa  
7 Cobo<sup>6,\*</sup>, Robert Castelo<sup>2,9,\*</sup>

8  
9 <sup>1</sup>Department of Pediatrics, Hospital de Figueres, Figueres, Spain; <sup>2</sup>Department of Experimental and  
10 Health Sciences, Universitat Pompeu Fabra (UPF), Barcelona, Spain; <sup>3</sup>Genomics Core Facility,  
11 Department of Experimental and Health Sciences, Universitat Pompeu Fabra (UPF), Barcelona, Spain;  
12 <sup>4</sup>Proteomics Unit, Centre de Regulació Genòmica (CRG), Barcelona, Spain; <sup>5</sup>Barcelona Institute of  
13 Science and Technology (BIST), Barcelona, Spain; <sup>6</sup>Hospital Clínic, Institut d'Investigacions Biomèdiques  
14 August Pi i Sunyer (IDIBAPS), University of Barcelona, Centre for Biomedical Research on Rare  
15 Diseases (CIBER-ER), Barcelona, Spain; <sup>7</sup>Hospital Sant Joan de Déu, Barcelona, Spain; <sup>8</sup>Department of  
16 Pathology, Hospital Clínic, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), University  
17 of Barcelona, Barcelona, Spain; <sup>9</sup>Research Programme on Biomedical Informatics, Institut Hospital del  
18 Mar d'Investigacions Mèdiques (IMIM), Barcelona, Spain; \*Correspondence: Robert Castelo  
19 ([robert.castelo@upf.edu](mailto:robert.castelo@upf.edu)), Teresa Cobo ([tcobo@clinic.cat](mailto:tcobo@clinic.cat))

20  
21 Word count (excluding Abstract, Methods, References and Figure and Table Legends): 3,638

22  
23

## 2 **Abstract**

3 The fetal inflammatory response (FIR) increases the risk of perinatal brain injury, particularly in  
4 extremely low gestational age newborns (ELGANs, < 28 weeks of gestation). One of the  
5 mechanisms contributing to such a risk is a postnatal intermittent or sustained systemic  
6 inflammation (ISSI) following FIR. The link between prenatal and postnatal systemic  
7 inflammation is supported by the presence of well-established inflammatory biomarkers in the  
8 umbilical cord and peripheral blood. However, the extent of molecular changes contributing to  
9 this association is unknown. Using RNA sequencing and mass spectrometry proteomics, we  
10 profiled the transcriptome and proteome of archived neonatal dried blood spot (DBS) specimens  
11 from 21 ELGANs. Comparing FIR-affected and unaffected ELGANs, we identified 783 gene and  
12 27 protein expression changes of 50% magnitude or more, and an experiment-wide significance  
13 level below 5% false discovery rate. These expression changes confirm the robust postnatal  
14 activation of the innate immune system in FIR-affected ELGANs and reveal an impairment of  
15 their adaptive immunity. In turn, the altered pathways provide clues about the molecular  
16 mechanisms triggering ISSI after FIR, and the onset of perinatal brain injury.

## 17 **Keywords**

18 Premature birth; fetal inflammatory response; perinatal brain injury; transcriptomics; proteomics

## 19 **Introduction**

20 Intraamniotic infection (IAI), one of the main causes of preterm birth, can lead to a maternal  
21 (MIR) and a fetal (FIR) inflammatory response, which can be detected in placental tissues<sup>1</sup>  
22 (Supp. Table S1). FIR can cause damage to fetal organs presumably by triggering a systemic  
23 inflammation, leading to distant organ injury and increasing the risk of multiple adverse neonatal  
24 outcomes, such as perinatal brain damage<sup>2,3</sup>. Understanding the molecular dynamics of FIR and  
25 its downstream alterations before and after birth is thus crucial to developing diagnostic and  
26 therapeutic strategies that improve the clinical outcome of extreme preterm birth.

27  
28 In a previous work, we found that FIR triggers a broad and complex transcriptional response in  
29 umbilical cord (UC) tissue<sup>4</sup>, which others have shown to correlate with cognitive impairment at  
30 10 years of age<sup>5</sup>. The molecular and cellular alterations induced by FIR can persist after birth<sup>6</sup>  
31 and interact with postnatal inflammation-initiating illnesses such as bacteremia, mechanical  
32 ventilation or necrotizing enterocolitis<sup>7</sup>. This interaction ultimately leads to what is known as an

2 intermittent or sustained systemic inflammation (ISSI)<sup>8,9</sup>, which is in turn strongly associated  
3 with perinatal brain injury and developmental disorders in ELGANs<sup>9</sup> (Extremely Low Gestational  
4 Age Newborn, i.e. birth before the 28 week of gestation). The joint presence of placental  
5 inflammation and ISSI appears to increase the risk of perinatal brain injury, highlighting the  
6 important contribution of postnatal systemic inflammation to that risk<sup>10,11</sup>. However, little is  
7 known about the extent of postnatal molecular changes associated with FIR, which can  
8 contribute to understanding the link between FIR and ISSI and consequently provide an array of  
9 candidate biomarkers and therapeutic targets for perinatal brain injury.

10

11 Here, we analyze the transcriptome and proteome of archived neonatal DBSs from ELGANs  
12 and identify RNA and protein expression changes in whole blood that are significantly  
13 associated with FIR. These molecular changes provide, with an unprecedented level of  
14 resolution, a snapshot of the pathways participating in ISSI as a result of the exposure to FIR.

## 15 **Results**

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

17 The samples analyzed in this study were from 21 ELGANs (8 females and 13 males), who met  
18 the eligibility criteria (see Methods). Histological acute chorioamnionitis was diagnosed in 15  
19 cases. MIR was in stage I in two cases, stage II in one case, and stage III in 12 cases. Ten of  
20 the 12 cases with stage III MIR showed FIR: four in stage I, five stage II, and one stage III  
21 (Supp. Fig. S1). Other clinical characteristics and outcomes are presented in Table 1,  
22 separately for those with and without FIR. For instance, FIR-affected ELGANs presented a  
23 higher frequency of microbial invasion of the amniotic cavity (89% vs 29%, two-tailed Fisher's  
24 exact test  $P=0.035$ ) than unaffected ones, and had significantly higher maternal blood C-  
25 reactive protein (CRP) levels at admission ( $t$ -test  $P=0.041$ ), as well as significantly higher ( $t$ -test  
26  $P=0.034$ ) and lower ( $t$ -test  $P=0.018$ ) amniotic fluid (AF) IL-6 and glucose levels, respectively.  
27 Our data are compatible with no differences in neonatal morbidities between FIR-affected and  
28 unaffected ELGANs, although a higher frequency of neonatal sepsis and intraventricular  
29 hemorrhage was observed in FIR cases. Mean gestational age (GA) and birth weight were  
30 similar between the two groups, averaging 26 weeks and 860 g, respectively. Such an  
31 extremely low birth weight, with a lowest value of 580 g and up to 8 neonates under 750 g,  
32 precludes a standard blood extraction after birth for the purpose of molecular profiling in  
33 research. On the other hand, UC blood is not an alternative either, because is not

2 representative of postnatal immunity<sup>12</sup>. For this reason, neonatal DBS samples archived at the  
 3 newborn screening program constitute a minimally invasive option to profile the transcriptome  
 4 and proteome in peripheral blood from ELGANs.

5

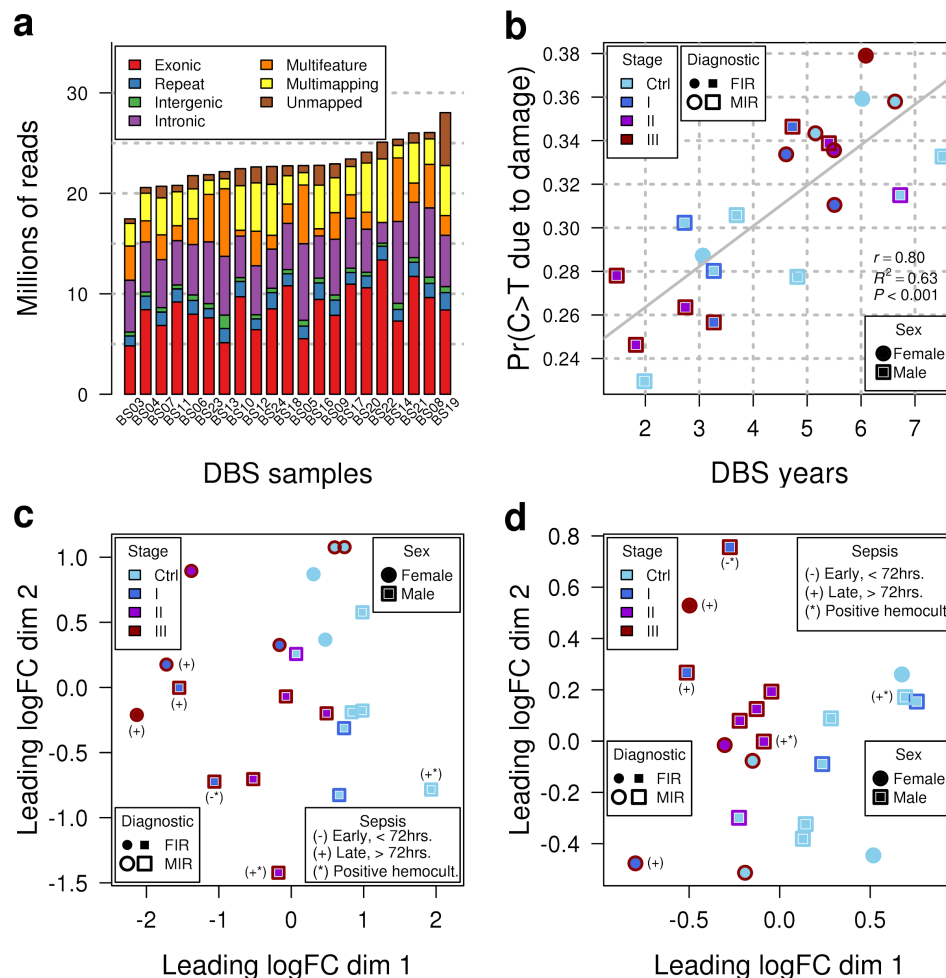
6 **Table 1 | Clinical characteristics of ELGANs**

Clinical characteristics	No FIR (n=11)	FIR (n=10)	P value
<i>Baseline characteristics of pregnancy and delivery</i>			
Cesarean section: n, (%)	8 (72.7)	3 (30.0)	0.086
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), mean±SD	2.3 ± 3.4	5.8 ± 3.8	0.041
<i>Biomarkers of intra-amniotic infection and inflammation</i>			
MIAC: n, (%)	2 (28.6) 4 missing obs.	8 (88.9) 1 missing obs.	0.035
Clinical chorioamnionitis: n, (%)	0 (0.0)	3 (30.0)	0.090
Histological chorioamnionitis: n, (%)	5 (45.5)	10 (100.0)	0.012
Concentration of IL-6 in AF: (pg/mL), (Q1, Q3)	21,153.5 (6,695.50, 35,975.75) 5 missing obs.	112,271.5 (34,968.07), 162,866.25) 2 missing obs.	0.034
Concentration of glucose: (mg/dL), (Q1, Q3)	23.6 (10, 37) 4 missing obs.	2.4 (0, 2) 1 missing obs.	0.018
<i>Baseline characteristics of infants</i>			
Male: n, (%)	7 (63.6)	6 (60.0)	1.000
Birth weight: (g), mean±SD	841 ± 138	880 ± 159	0.552
Respiratory distress syndrome: n, (%)	5 (45.5)	5 (50.0)	1.000
Patent ductus arteriosus: n, (%)	6 (54.5)	6 (60.0)	1.000
Intraventricular hemorrhage: n, (%)	1 (9.1)	5 (50.0)	0.063
White matter disease: n, (%)	1 (9.1)	2 (20.0)	0.586
Retinopathy of prematurity: n, (%)	5 (45.5)	3 (37.5)	1.000

		2 missing obs.	
Necrotizing enterocolitis: n, (%)	1 (9.1)	0 (0.0)	1.000
Sepsis: n, (%)	1 (9.1)	5 (50.0)	0.063
Positive blood culture: n, (%)	1 (9.1)	2 (20.0)	0.586
GA 23-24 completed weeks: n, (%)	2 (18.2)	1 (10.0)	1.000
GA 25-26 completed weeks: n, (%)	1 (9.1)	5 (50.0)	0.063
GA 27 completed weeks: n, (%)	8 (72.7)	4 (40.0)	0.198
DBS heel sampling days > 3: n, (%)	4 (36.4)	5 (50.0)	0.678
Maximum CRP within first 8 postnatal days: (md/dL), mean±SD	1.0 ± 2.2	3.5 ± 5.3	0.187
WBC count at birth: (10 <sup>9</sup> /L), 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), mean±SD	13.2 ± 5.2	44.1 ± 13.4	< 0.001
Maximum ANC within first 8 postnatal days: (10 <sup>9</sup> /L), mean±SD	6.2 ± 4.3	29.0 ± 10.8	< 0.001
ALC at birth: (10 <sup>9</sup> /L), 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), 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

2 CRP: C-reactive protein. AF: amniotic fluid. MIAC: microbial invasion of the amniotic cavity defined by a positive AF culture. WBC:  
3 white blood cell. ANC: absolute neutrophil count. ALC: absolute lymphocyte count. Intraventricular hemorrhage: any grade between  
4 I to IV according Papile et al (1978)<sup>13</sup> grading system. Sepsis: confirmed or suspected bacterial blood infection, which is classified  
5 as early if within 72 hours after birth and classified as late if after the third postnatal day. Categorical characteristics were tested with  
6 respect to FIR status using a two-tailed Fisher's exact test. Continuous characteristics were tested with respect to FIR status using a  
7 *t*-test.

8



2  
 3 **Figure 1:** Transcriptomics and proteomics profiling of archived DBSs. (a) Sequencing depth per sample,  
 4 broken by origin of sequence reads. (b) Probability of a C>T deamination due to damage, as function of  
 5 DBS age in years. (c) Sample differences in terms of fold-changes between transcriptomic samples in  
 6 base-2 logarithmic scale (logFC), projected along the x and y-axes. Distance between points is  
 7 proportional to the dissimilarity between samples. (d) Same as (c) but calculated from proteomic samples.

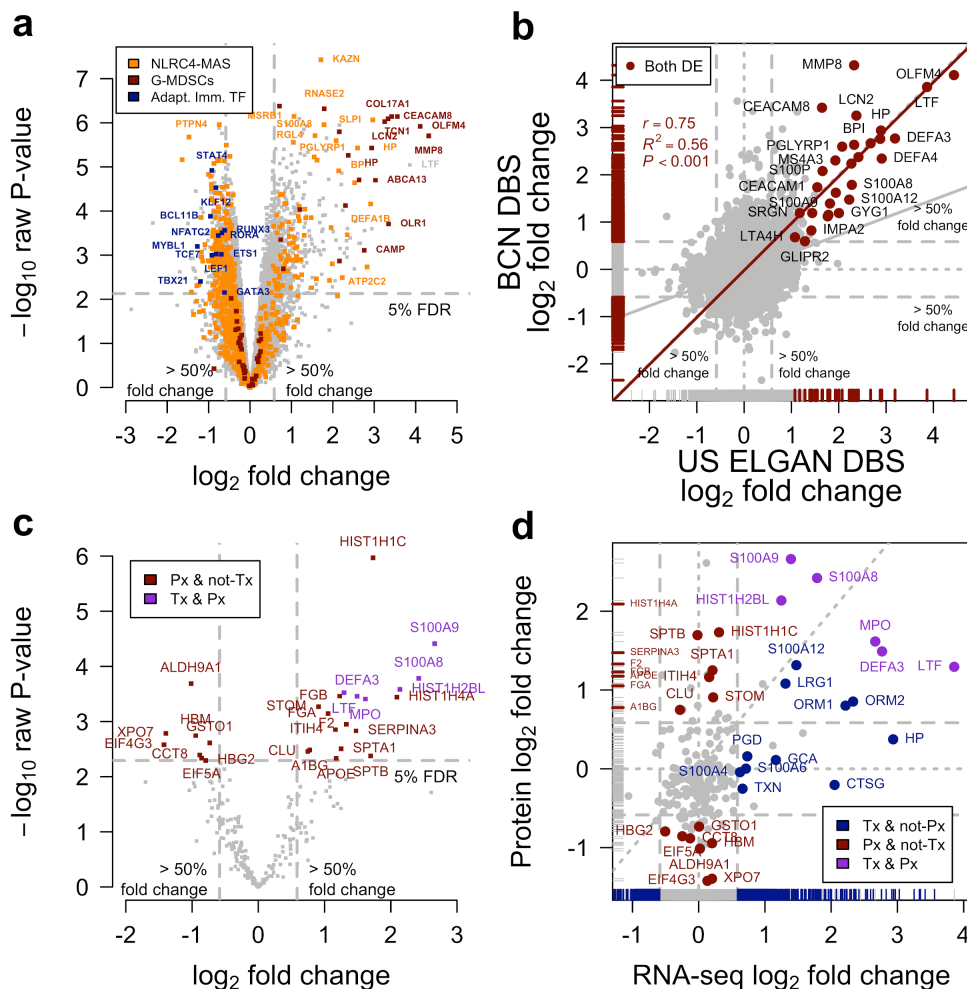
8  
 9 Total RNA sequencing of the 21 archived DBSs produced more than 500 million paired-end  
 10 reads, which we transformed into a table of 25,221 gene-level summarized count expression  
 11 profiles. Nearly half of these genes (11,279) showed reliable levels of expression in at least 6 of  
 12 the 21 samples and were those selected for further analysis (see Methods). Figure 1a shows  
 13 the distribution of sequencing depth across different genomic elements for each sample, where  
 14 only about  $\frac{1}{3}$  of the total sequencing depth derives from exonic reads. DBS samples in this  
 15 cohort were stored at room temperature for one to seven years until RNA extraction and  
 16 sequencing. Over this period of time, degradation and fragmentation processes affect RNA  
 17 integrity and we thus investigated whether, as observed in ancient DNA samples<sup>14</sup>, there is a

2 detectable damage to the RNA due to cytosine deamination on such a short timescale (see  
3 Methods). Indeed, we found a significant linear relationship ( $r=0.8$ ,  $P < 0.001$ ) between the  
4 probability of a C>T substitution due to damage and the elapsed time between DBS extraction  
5 and RNA sequencing (Fig. 1b). Additionally, mass-spectrometry proteomics on DBSs produced  
6 649 quantified protein expression profiles (Supplementary Data File 1), from which we selected  
7 245 as being reliably expressed in at least 6 of the 21 samples (see Methods).

8 Sample-level changes of transcriptomics data, projected in two dimensions (Fig. 1c), show that  
9 gene-level expression profiles derived from exonic reads capture the differences in placental  
10 inflammation throughout the different stages of MIR and FIR. Analogously, DBS protein  
11 expression profiles also show a clear separation between infants with and without MIR and FIR  
12 (Fig. 1d). In summary, transcriptomics and proteomics of archived DBS samples can produce a  
13 large catalog of molecular measurements that capture clinical features of the perinatal period.

#### 14 **Archived neonatal DBSs contain gene and protein expression changes associated with** 15 **FIR**

16 The differential expression analysis of RNA-seq transcriptomic profiles between FIR and non-  
17 FIR exposed ELGANs identified 783 differentially expressed (DE) genes under an experiment-  
18 wide significance threshold of 5% false discovery rate (FDR) and a minimum 50%-fold change  
19 (Fig. 2a, Supp. Table S2). We replicated similar magnitude changes from DE genes on a  
20 logarithmic scale ( $r=0.75$ ,  $R^2=0.56$ ,  $P < 0.001$ ) in the analogous differential expression analysis  
21 of a DBS gene expression dataset<sup>15</sup> from the US ELGAN cohort (Fig. 2b, Supp. Table S3 and  
22 Supp. Fig. S2 to S5), despite differences in sample preparation and gene expression profiling  
23 technology (see Methods). The hierarchical clustering of the 21 transcriptomic samples, using  
24 the 783 DE genes, leads to a perfect separation of the neonates who were affected by FIR, from  
25 those that were unaffected (Fig. 3).



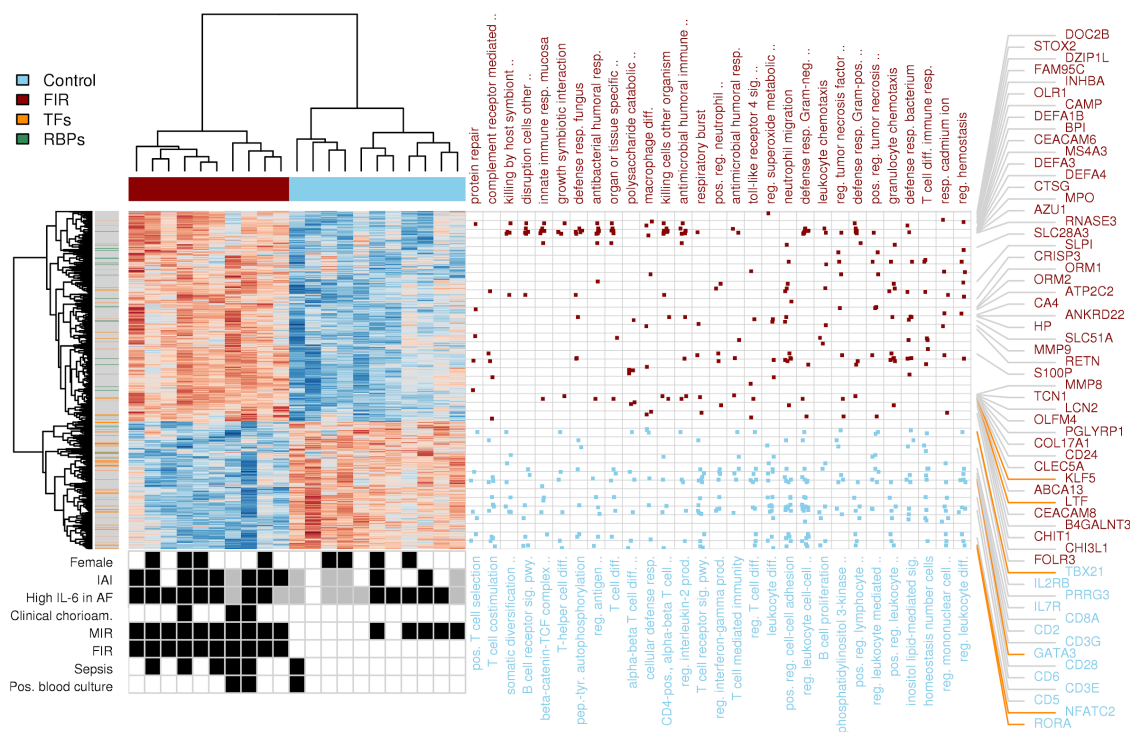
2  
 3 **Figure 2:** Differential expression analysis between FIR-affected and unaffected ELGANS, of  
 4 transcriptomic and proteomic profiles from archived DBSs. (a) Volcano plot of transcriptomics data. Raw  
 5  $P$  values in negative logarithmic scale on the y-axis as function of the  $\log_2$ -fold change on the x-axis.  
 6 Orange dots highlight genes from a NLRC4-MAS signature<sup>16</sup>, red dots highlight genes from a G-MDSCs  
 7 signature<sup>17</sup> and blue dots highlight transcription factor coding genes involved in regulating adaptive  
 8 immunity. (b) Fold changes of this study on the y-axis as function of fold changes in data from the US  
 9 ELGAN cohort in the x-axis. (c) Volcano plot of proteomics data. (d) Fold changes from proteomics data  
 10 on the y-axis as function of fold changes from transcriptomics data on the x-axis.

11  
 12 A further question is whether there are RNA changes in peripheral blood across the different  
 13 stages of MIR and FIR. Indeed, we found 58 genes with  $FDR < 5\%$ , for which a combination of  
 14 MIR and FIR stages, reflecting an increasing exposure to prenatal inflammation, led to a  
 15 minimum increase of 50% in gene expression levels for each additional combined inflammatory  
 16 stage (Supplementary Figure S6 and Table S4).

17



2 The differential expression analysis of proteomic profiles between FIR-exposed neonates and  
 3 those who were unexposed identified 27 DE proteins at 5% FDR with a minimum 50%-fold  
 4 change (Fig. 2c, Supp. Table S5). Figure 2d shows a comparison between log2-fold changes of  
 5 the transcriptomics and proteomics analyses. Six of the 27 DE proteins are encoded by genes  
 6 that are also DE by RNA-seq, while for 14 of them, their corresponding genes do not show fold-  
 7 changes above 50% in the transcriptomics data. The remaining seven DE proteins do not have  
 8 minimum levels of expression in the RNA-seq assay. Among them, F2, FGA, FGB and A1BG  
 9 are secreted from the liver into the blood, and this may also be the case for SERPINA3 and  
 10 APOE, which have a biased expression towards liver. Overall, these results show that in  
 11 peripheral blood from ELGANs there are sizeable numbers of transcriptomic and proteomic  
 12 changes associated with FIR that can be detected in archived DBS.  
 13



14  
 15 **Figure 3:** FIR gene expression signature and functional enrichment analysis. Heatmap of expression  
 16 values for 783 DE genes with FDR < 5% and minimum 1.5-fold change between FIR-affected and  
 17 unaffected ELGANs, obtained after removing FIR-unrelated variability. Dendrograms on the x and y-axes  
 18 represent the hierarchical clustering of samples and genes, respectively. Leaves on the gene dendrogram  
 19 are color coded according to whether genes encode for transcription factors (orange) or RNA-binding  
 20 proteins (green), while those on the sample dendrogram are color coded according to FIR status, where  
 21 red indicates samples derived from FIR-affected infants and blue unaffected. The right dot-matrix  
 22 represents DE genes (y-axis) belonging to GO terms (x-axis) significantly enriched (FDR < 10% and

2 odds-ratio -OR > 1.5) by upregulated (top) and downregulated (bottom) DE genes. The bottom-left dot-  
3 matrix represents the presence or absence of phenotypes, with grey indicating missing data.

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

5 Protein biomarkers have helped in the discovery of a postnatal systemic inflammatory hit  
6 following FIR<sup>6</sup>. However, the extent of such molecular changes was previously unknown. Our  
7 transcriptomic data from DBS samples shows the upregulation in FIR of many genes involved in  
8 the innate immune response (Fig. 2a; Supp. Table S2). Such genes include those coding for  
9 pattern recognition receptors (PRRs, e.g., *NLRC4*, *CLEC4E*, *FPR1*), inflammatory transcription  
10 factors (e.g., *CEBPD*, *KLF5*), acute phase response proteins (*HP*, *ORM1*), as well as multiple  
11 inflammatory mediators including cytokines (e.g., *IL18*, *IL1R1*, *RETN*), calgranulins (S100A8,  
12 S100A9, S100A12), chemokines (e.g., *CXCR1*, *CXCL1*), complement (e.g., *C5AR2*, *C3AR1*),  
13 lipid mediators (e.g., *LTA4H*, *ALOX5AP*, *CYP4F3*), proteolytic enzymes (e.g., *MMP8*, *MMP9*)  
14 and adhesion molecules (e.g., *ICAM3*, *ITGAM*). Our proteomic data from DBS samples (Fig. 2c,  
15 Supp. Table S5) also shows protein overexpression in FIR of diverse inflammatory mediators of  
16 the innate immune system, such as the S100A8, S100A9, LTF, MPO and DEFA3 proteins.

17  
18 A functional enrichment analysis using the Gene Ontology (GO) database (Fig. 3) shows GO  
19 terms associated with the innate immune system, such as *innate immune response in mucosa*  
20 (OR=17.7) and *defense response to fungus* (OR=14.3), which were significantly enriched (FDR  
21 < 10%) in upregulated genes (Supp. Table S6; see Methods). Moreover, we also observed a  
22 significant enrichment of DE genes in curated gene sets of the innate immune response, such  
23 as the InnateDB<sup>18</sup> (one-tailed Fisher's exact P < 2.2 × 10<sup>-16</sup>, OR=3.2). All these gene and  
24 protein expression changes constitute the most comprehensive description to date of the  
25 persistent postnatal activation of the innate immune system in ELGANs affected by FIR.

#### 26 **FIR is associated with NLRC4 inflammasome activation**

27 The activation of the innate immune system involves the assembly of different multiprotein  
28 complexes known as inflammasomes, which are categorized according to their constituent  
29 PRRs. After detection of microbial and endogenous danger signals by PRRs, inflammasomes  
30 induce, via activation of caspase-1 (CASP1), the maturation and release of IL-1 $\beta$  and IL-18  
31 cytokines, bacterial clearance and pyroptosis. The release of interleukin-1 family cytokines such  
32 as IL-1 $\beta$  and IL-18 has potent effects on neutrophil activation and recruitment<sup>19</sup>. The NLRC4

2 inflammasome belongs to the intracellular cytosolic PRR family of NOD-like receptors and  
3 detects the cytosolic presence of bacterial components.

4  
5 We found that FIR upregulates the expression of the NLRC4 inflammasome and IL-1 family-  
6 related protein-coding genes (e.g., *NLRC4*, *PYCARD*, *IL18*, *IL1R1*). We also found a significant  
7 overlap (Fig. 2a, one-tailed Fisher's exact  $P < 2.2 \times 10^{-16}$ , OR=10.4) of DE genes with a  
8 transcriptional signature of the NLRC4-infantile-onset macrophage activation syndrome  
9 (MAS)<sup>16</sup>. Consistent with these results, we observed the upregulation of numerous genes  
10 involved in neutrophil biology and function (Supp. Table S7), significantly enriching (FDR <  
11 10%, OR > 1.5) GO terms related to neutrophil activation<sup>20</sup>, such as *neutrophil migration*  
12 (OR=6.3) or *positive regulation of neutrophil chemotaxis* (OR=7.5); see Fig. 3 and  
13 Supplementary Table S6. These findings suggest that FIR leads to an initial proinflammatory-  
14 driven response (*IL-6*, *IL-1b*, *CXCL8*), followed by a posterior neonatal robust activation of the  
15 IL-18/IL-1 axis through the NLRC4-inflammasome that can induce a persistent, systemic  
16 neutrophil activation.

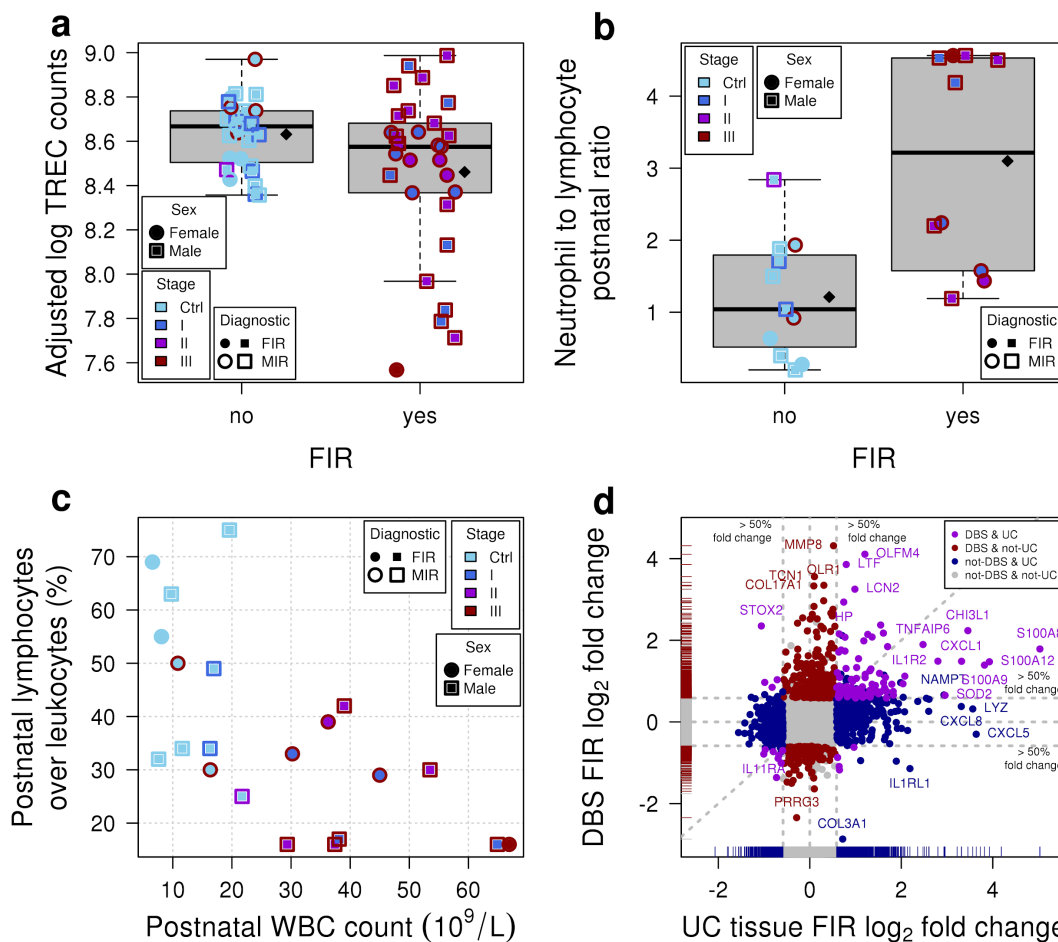
### 17 **FIR is associated with the impairment of adaptive immunity and thymic function through** 18 **the expansion of granulocytic myeloid-derived suppressor cells**

19 The mechanism by which antenatal FIR leads to a postnatal systemic inflammation is unknown.  
20 Among several plausible alternatives, Dammann and Leviton (2014)<sup>8</sup> suggested an alteration in  
21 the number of T cells and their regulation. A closer look at the hierarchical clustering of genes in  
22 Figure 3 reveals an enrichment of downregulated transcription-factor coding genes (one-tailed  
23 Fisher's exact test  $P < 0.001$ , OR=2.9, Supp. Table S8) in FIR-affected ELGANs. Many of these  
24 genes are regulators of T cell development and activation, such as *TBX21*, *GATA3* and *RORA*,  
25 respectively involved in T-helper(h) fate of CD4+ T cells in Th1, Th2 and Th17 lineages (Fig.  
26 2a). More generally, we found the downregulation of key genes associated with the IL-7 and IL-  
27 2 signaling pathways (e.g., *IL7R*, *BCL2*, *PIK3AP1*, *IL2RB*, *CD3E*, *CD3G*) involved in T cell  
28 development (Supp. Tables S2 and S9). The TCR signaling pathway (e.g., *CD3E*, *CD3G*, *LCK*,  
29 *FYN*, *ZAP70*, *CD247*) and T cell differentiation co-stimulatory receptor genes (e.g., *CD28*,  
30 *ICOS*, *SLAMF1*) were also downregulated<sup>21</sup>. Likewise, the inhibitory and stimulatory genes  
31 *KLRC1*, *KLRC3*, *KLRC4* that belong to the natural killer (NK) cell lectin-like receptors of the NK  
32 gene complex, and genes involved in NK cell cytotoxicity, such as *PRF1*, *SH2D1B* genes, were  
33 also downregulated in FIR-affected ELGANs<sup>21</sup>. Interestingly, we observed a significant overlap

2 of downregulated genes with an NLRC4-infantile-onset MAS signature (one-tailed Fisher's exact  
3  $P < 2.2 \times 10^{-16}$ ,  $OR=21.5$ )<sup>16</sup>, which is associated with cytotoxic and regulatory T cell dysfunction.

4  
5 Recent thymic emigrants consist of phenotypically and functionally immature T cells that have  
6 recently egressed from the thymus, which predominate in the fetus and newborns. These cells  
7 can be identified by their content of T cell receptor excision circles (TRECs)<sup>22</sup>, which are DNA  
8 fragments excised during the rearrangement of segments of T cell receptor (TCR) genes in the  
9 thymus. TRECs are used as a measure of the number of T lymphocytes in newborns and,  
10 indirectly, as an indicator of a possible thymus dysfunction. We found downregulation of the  
11 genes coding for the PTK7 and S1PR1 proteins in FIR-affected ELGANs. PTK7 is a cell marker  
12 of CD4+ T cell recent thymic emigrants used for assessing thymic activity<sup>23</sup>, while the S1PR1  
13 receptor expression in CD4+ T cells is required for their survival and thymic emigration<sup>24</sup>.  
14 Consistent with these results, FIR-affected ELGANs had significantly lower levels of TRECs in  
15 the analyzed DBSs ( $P=0.0432$ ; Fig. 4a, see Methods), a significantly higher postnatal neutrophil  
16 to lymphocyte ratio ( $P=0.002$ ; Fig. 4b), and a significantly lower percent of all leukocytes that  
17 are lymphocytes at birth ( $P=0.012$ ; Table 1) and during the first postnatal week<sup>25</sup> ( $P=0.003$ ; Fig.  
18 4c).

19  
20 Myeloid-derived suppressor cells (MDSCs) develop under pathological conditions from myeloid  
21 precursors. According to their phenotype, MDSCs can be classified into monocytic(M)-MDSCs  
22 or granulocytic(G)-MDSCs. Postnatally, there is a high level of circulating MDSCs,  
23 predominantly G-MDSCs, which correlates with inflammatory markers of perinatal infection<sup>26,27</sup>.  
24 Neonatal MDSCs display reduced apoptosis and immunosuppressive activity after bacterial  
25 infection<sup>28</sup>. Indeed, these cells are potent inhibitors of adaptive immune responses, particularly  
26 of T and NK cells, by different mechanisms, including the production of reactive oxygen species  
27 (ROS)<sup>29-31</sup>, prostaglandin E2 (PGE2), and the depletion of L-arginine by the arginase 1 enzyme  
28 (ARG1)<sup>32,33</sup>. In our data, we found that the *ARG1* gene was 2.5-fold upregulated in FIR-affected  
29 ELGANs, as well as the *OLR1* gene, encoding the lectin-type oxidized LDL receptor 1 (LOX1), a  
30 marker of G-MDSCs<sup>17</sup>, which was over 10-fold upregulated. We also found a significant overlap  
31 (one-tailed Fisher's exact  $P < 2.2 \times 10^{-16}$ ,  $OR=26.4$ ) between FIR upregulated genes and a G-  
32 MDSC expression signature<sup>17</sup> (Fig. 2a).



2  
 3 **Figure 4:** Association of FIR with clinical and molecular signatures. (a) TREC corrected fluorescent  
 4 counts on logarithmic scale by FIR status, after adjusting for sex, maximum lymphocyte count during the  
 5 first postnatal week and batch. Diamonds indicate mean values. There are multiple measurements per  
 6 newborn. (b) Postnatal ratio of neutrophil absolute count to absolute lymphocyte count, during first  
 7 postnatal week, by FIR status. Diamonds indicate mean values. (c) Percent of all leukocytes that are  
 8 lymphocytes during first postnatal week on the y-axis as function of the postnatal white blood cell count  
 9 on the x-axis. (d) Fold changes in this study on the y-axis as function of FIR fold changes in UC tissue on  
 10 the x-axis.

11  
 12 One of the mechanisms by which G-MDSCs suppress T cell responses is the release of ROS  
 13 molecules. We found an upregulation of genes coding for proteins essential to ROS production  
 14 such as *CYBB*, *NCF1*, *NCF4*, *MPO*, *RAB27A*, as well as genes coding for antioxidant enzymes  
 15 (e.g., *SOD2*, *TXN* and the Msr family -*MsrA*, *MsrB1*, *MsrB2*, *MsrB3*). Genes encoding Msr  
 16 proteins reduce protein methionine oxidation, a post-translational alteration that correlates with  
 17 oxidative stress<sup>34</sup>. In summary, our data suggest the expansion of G-MDSCs as a mechanism

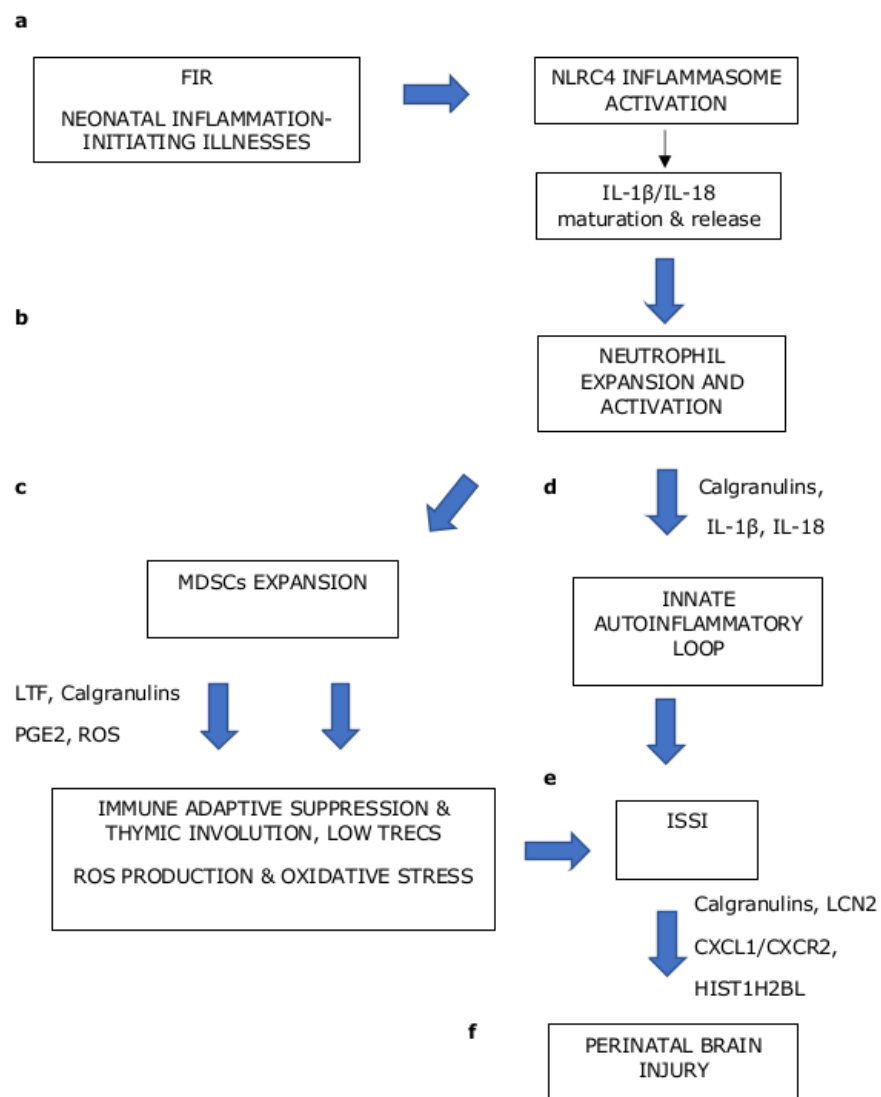
2 contributing to the impairment of adaptive immunity in ELGANs affected by FIR, and  
3 consequently, to a postnatal systemic inflammation sustained over time after birth.

#### 4 **Discussion**

5 We have described the largest catalog to date of postnatal transcriptomic and proteomic  
6 changes associated with FIR in archived DBSs from ELGANs. This catalog demonstrates that  
7 archived neonatal DBSs constitute a valuable source of genome-wide molecular information  
8 about the perinatal period. More importantly, we have shown that molecular changes in this  
9 catalog provide clues to the underlying relationships between FIR and ISSI.

10  
11 Our data provide molecular evidence of the postnatal activation of NLRC4-inflammasome  
12 dependent mechanisms, which may contribute to the development of a postnatal innate immune  
13 response in FIR-affected ELGANs (Fig. 5a). The NLRC4-inflammasome hyperactivity can cause  
14 diverse auto-inflammatory diseases, such as the infantile-onset MAS syndrome and  
15 enterocolitis. These disorders are characterized by a huge elevation of IL-18 in the blood<sup>16</sup>.  
16 Recently, this protein has been proposed as a therapeutic target in neonatal sepsis<sup>35</sup>, and in  
17 fact, we observed a 2-fold overexpression of the *IL18* gene in FIR-affected infants. In preterm  
18 newborns, Liang et. al (2017)<sup>36</sup> reported the case of a 28-week preterm neonate with a *de novo*  
19 gain-of-function mutation in the *NLRC4* gene, who died of a fatal syndrome of excessive  
20 immune activation likely to have begun in utero. NLRC4 dependent mechanisms can also  
21 contribute to brain injury induced by cerebral ischemia, as has recently been described in a  
22 rodent-based model of stroke<sup>37</sup>. Interestingly, caffeine therapy for the apnea of prematurity,  
23 which reduces the incidence of moderate to severe neurodevelopmental disabilities in very low  
24 birth weight infants<sup>38</sup>, can also reduce NLRC4-inflammasome activation<sup>39</sup>.





2  
3 **Figure 5:** Model for the association between FIR, ISSI and perinatal brain injury. (a) FIR and postnatal  
4 diseases can induce the postnatal activation of NLR4-inflammasome dependent mechanisms in innate  
5 immune cells that trigger the release of proinflammatory cytokines IL-1 $\beta$  and IL-18. (b) These cytokines  
6 have a major role in neutrophil expansion and activation. (c) The neutrophil proliferation under  
7 pathological conditions can induce the expansion of myeloid-derived suppressor cells (MDSCs). These  
8 cells are potent suppressors of T and NK cells and strong producers of radical oxygen species (ROS). (d)  
9 Robust production of calgranulins and IL-1 $\beta$ / IL-18 cytokines can lead to an innate autoinflammatory loop  
10 by autocrine feedback-loops. (e) The innate autoinflammatory loop, the immune adaptive suppression  
11 and oxidative stress, can lead to ISSI. (f) ISSI can cause the infiltration of innate immune cells into the  
12 brain, such as neutrophils, induced by diverse inflammatory mediators, such as calgranulins, leading to  
13 perinatal brain injury in preterm newborns.

14

2 The release of the proinflammatory cytokines IL-18 and IL-1 $\beta$  after NLRC4 inflammasome  
3 activation induces the expansion, migration and activation of neutrophils (Fig. 5b). Molecular  
4 alterations in neonatal neutrophils, such as a reduced expression of Fas receptors, can lead to  
5 delayed neutrophil apoptosis<sup>40,41</sup> and the persistence of proinflammatory responses<sup>42</sup>. We found  
6 a significantly higher neutrophil to lymphocyte ratio in FIR-affected ELGANs (Fig. 4b), the  
7 overexpression of *S100A8*, *S100A9*, *IL18*, *MMP8* and *OLFM4*, and the underexpression of  
8 gasdermin B genes (Supp. Table S2). Calgranulins S100A9 and S100A8, and IL-18 can inhibit  
9 neutrophil apoptosis<sup>43,44</sup> and neutrophils can evade pyroptosis, which is mediated by gasdermin  
10 family proteins<sup>45</sup>. Furthermore, the robust production of calgranulins and the activation of the IL-  
11 1/IL-18 axis can trigger a neutrophil-associated inflammation-boosting loop binding to PRRs and  
12 triggering downstream inflammatory signaling pathways, just as in auto-inflammatory  
13 disorders<sup>46</sup>. Among children with septic shock, elevated *MMP8* and *OLFM4* blood gene  
14 expression levels, are associated with higher rates of mortality and organ failure<sup>47,48</sup>. Similarly,  
15 other studies have also shown that preterm neonates with abnormal neuroimaging studies had  
16 increased blood neutrophil activation markers<sup>49,50</sup>.

17  
18 The analysis of downregulated genes supports the inhibition of adaptive immune responses in  
19 FIR (Fig. 1a and 3), which is consistent with the relationships we see between TRECs and FIR  
20 (Fig. 4a), as well as the lymphocyte percent among ELGANs with and without FIR and MIR (Fig.  
21 4c). The reduction of TRECs and the downregulation of the *PTK7* and *S1PR1* genes suggest a  
22 thymic involution associated with FIR<sup>51</sup>, which may lead to a relative increase in  
23 proinflammatory CD31-CD4+ T cells producing TNF<sup>52</sup>. Perhaps that is why thymic involution has  
24 also been associated with increased risk of brain damage in ELGANs<sup>53</sup>. On the other hand, in  
25 our earlier investigation<sup>4</sup> of UC tissue from ELGANs, we found multiple upregulated genes in  
26 FIR that code for proteins involved in the expansion of MDSCs (Fig. 4d). In the present study,  
27 we found a significant overlap between FIR DE genes and a previously published G-MDSC  
28 signature<sup>17</sup> (Fig. 5c).

29  
30 In newborns, the LTF protein appears to promote the secretion of S100A9<sup>54</sup>, which can induce  
31 the production of PGE2, thereby repressing the expression of LCK and ZAP70 proteins in CD4+  
32 T cells, and impairing their activation<sup>30,31,55</sup>. Overall, our results suggest that FIR may induce an  
33 initial prenatal pro-inflammatory-driven response (IL-6, IL1-b, CXCL8) evident in UC tissue and  
34 UC blood, followed by neonatal innate immune activation mediated by the NLRC4-  
35 inflammasome. As a result of these events, FIR-affected ELGANs display a robust neutrophil



2 and G-MDSC expansion during the first postnatal days, which can lead to adaptive immune  
3 suppression and the onset of an auto-inflammatory loop, contributing to ISSI and consequently  
4 increasing the risk of perinatal brain injury (Fig. 5d).

5  
6 An important feature of G-MDSCs is the production of ROS molecules, which can lead to  
7 oxidative stress and damage. Moreover, preterm newborns have a reduced antioxidant capacity  
8 that increases the risk of ROS-induced damage, which is an important factor in numerous  
9 diseases in premature infants, such as perinatal brain damage<sup>56</sup>. In fact, ELGANs are born with  
10 deficient levels of selenium<sup>57</sup>, a cofactor for the MsrB and thioredoxin enzyme activity required  
11 for proper antioxidant function. We found a 10-fold overexpression in FIR of both the *OLR1*  
12 gene, which encodes LOX1, a marker of G-MDSCs<sup>17</sup>, and the gene coding for the heat shock  
13 protein HSPA1A. HSPA1A can activate LOX1 in UC neutrophils in response to in vitro exposure  
14 to microbial peptidoglycan<sup>58</sup>, leading to the downstream production of ROS. However, the LTF  
15 protein, which we found to be significantly overexpressed in both transcriptomics and  
16 proteomics data, can protect newborns from oxidative stress and damage by attenuating  
17 inflammatory responses and controlling oxidative cell injury induced by innate immune  
18 activation<sup>59</sup>.

19  
20 Using well-established biomarkers, it has been shown that antenatal and postnatal  
21 inflammation, the so-called two-inflammatory hit, increases the risk of perinatal brain damage in  
22 preterm newborns<sup>10</sup>. Here, we found the overexpression of multiple genes encoding proteins  
23 that can be associated with perinatal brain damage in FIR-exposed newborns (Fig. 5e), such as  
24 LCN2, CXCL1, CXCR2, S100A8 and S100A9, as well as histone-related genes (e.g.,  
25 HIST1H2BL). LCN2 is an iron-binding protein contained in neutrophil granules that is associated  
26 with neuroinflammation<sup>60</sup>. Using an animal model of chorioamnionitis, Yellowhair et al. (2018)<sup>61</sup>  
27 found that chorioamnionitis-induced fetal brain damage was associated with increased  
28 expression of CXCL1 and CXCR2 in placental and fetal brain, and with an elevated number of  
29 cerebral CXCR2+ neutrophils. With respect to calgranulins, it has been observed that the brain  
30 expression of S100A8 and S100A9 proteins was increased in the brain of patients dying of  
31 sepsis; specifically, S100A9 expression was required for the cerebral recruitment of neutrophils  
32 and microglia activation in an animal model of sepsis<sup>62</sup>. Finally, extracellular histones can act as  
33 damage-associated molecular patterns, such as calgranulins, activating inflammatory signaling  
34 pathways and inducing multiple organ-specific damage<sup>63</sup>. In an animal study, histone H1 was  
35 found to induce proinflammatory responses in microglia and astrocytes<sup>64,65</sup>. Remarkably, we

2 found that both the gene and protein expression of S100A9, S100A8 and HIST1H2BL, was  
3 significantly higher in FIR-affected ELGANs, while our proteomics data also revealed the  
4 significant overexpression of HIST1H4A and HIST1H1C (Fig. 2c,d). In summary, our findings  
5 provide new insights into the molecular mechanisms that trigger ISSI after FIR, which may help  
6 to find new diagnostic biomarkers and therapeutic targets for pathological conditions derived  
7 from extreme prematurity, such as perinatal brain injury.

## 8 **Methods**

### 9 **Study design**

10 We performed a retrospective chart review of extreme preterm births (< 28 weeks of  
11 gestation) at the *Hospital Clinic de Barcelona* admitted to the neonatal intensive care unit  
12 (NICU) between January 2010 and October 2016. Two hundred thirty-four cases born alive  
13 were found. From this initial cohort, we excluded newborns from multiple pregnancies, outborn  
14 deliveries and those who either died in the delivery room or within the first week after birth,  
15 presented major congenital anomalies, received a blood transfusion before collecting the DBS,  
16 had no histopathological examination of the placenta and umbilical cord, or whose DBS was  
17 collected later than 8 days after birth. We attempted to maximize the number of individuals with  
18 available biomarkers of intra-amniotic infection and inflammation, while recruiting a similar  
19 number of FIR and non-FIR cases within a total maximum of 25. A final number of n=21 infants  
20 met the eligibility criteria and their families provided written consent to use their DBS samples  
21 for this study. At the time of processing these DBS samples, they had been stored at room  
22 temperature by the Catalan newborn screening program for between 1 and 7 years. The study  
23 protocol was approved by the Institutional Review Board from the *Hospital Clínic de Barcelona*  
24 (September 16<sup>th</sup>, 2016; ref. HCB/2016/0713).

### 25 **Clinical outcomes**

26 GA at birth was defined according to the first-trimester ultrasound examination. Preterm birth  
27 was classified by the clinical presentation: preterm labor, preterm pre-labor rupture of  
28 membranes, cervical insufficiency, placenta abruption, and preeclampsia. The laboratory results  
29 of AF cultures (genital mycoplasma, aerobic and anaerobic) and IL-6 levels, as well as the  
30 mother's blood CRP, WBC and ANC at admission were also recorded; see Table 1. Clinical  
31 chorioamnionitis was defined following the criteria of Gibbs<sup>66</sup>. Slides from the placenta and  
32 umbilical cord were independently reviewed by two experienced pathologists blinded to

2 experimental results and previous diagnosis. Maternal (MIR) and fetal (FIR) reactions were  
3 staged and graded according to published criteria<sup>1,67</sup>. Cases read discrepantly by the two  
4 pathologists were simultaneously re-reviewed on a two-headed microscope in order to reach a  
5 consensus (Supp. Table S1 and Supp. Fig. S1). The annotated neonatal morbidity included:  
6 respiratory distress syndrome (RDS); intraventricular hemorrhage (IVH); white matter disease  
7 (WMD); retinopathy of prematurity (ROP); sepsis; necrotizing enterocolitis (NEC); patent ductus  
8 arteriosus (PDA); and laboratory investigations at birth, and during the first 8 days after birth  
9 (CRP concentration, WBC and ANC). Hypothesis tests on clinical variables used in Table 1  
10 include two-tailed Fisher's exact test for categorical variables and *t*-tests for those that were  
11 continuous.

## 12 **Analysis of T cell receptor excision circles**

13 TREC data were obtained using the EnLite™ Neonatal TREC kit (PerkinElmer, Turku, Finland),  
14 which is a combination of PCR-based nucleic acid amplification and time-resolved fluorescence  
15 resonance energy transfer (TR-FRET) based detection. EnLite™ simultaneously detects two  
16 targets, TREC and beta-actin, the latter used as an internal control for monitoring specimen  
17 amplification in each test. The assay involves DNA elution from 1.5mm DBS punches,  
18 amplification and hybridization with target-sequence specific probes, and quantifies TREC levels  
19 by measuring probe fluorescence with the Victor Enlite™ fluorometer (PerkinElmer). A full  
20 calibration curve with blanks and three DBS calibrators was run in triplicate on each plate. A  
21 low-TREC control, a no-TREC control, a high-TREC control, and a blank paper disk (containing  
22 no blood), were used as quality controls for each plate. Raw fluorescence counts, measured at  
23 615 nm, 665 nm, and 780 nm, were processed by the EnLite™ workstation software to produce  
24 corrected fluorescence counts. TREC and beta-actin copies/μL were predicted from corrected  
25 fluorescent counts by the EnLite™ workstation software using an unweighted linear regression  
26 model on the ArcSinh transformation of the DBS calibrator copies/μL response.

27  
28 The values for the number of TREC copies/μL are sensitive to technical effects associated with  
29 the run in which they are obtained; for this reason, we generated technical replicates for the  
30 same individuals in different runs. Because of limited DBS availability, we could not produce  
31 more than one TREC read-out for all individuals and not all individuals with multiple readouts  
32 had the same number of replicated measurements. For this reason, we used a linear mixed-  
33 effects model to test the effect of FIR on the TREC read-out, using the newborn identifier as a  
34 random effect variable to account for replicated measurements across individuals. We used the

2 logarithm of the corrected fluorescent counts as TREC read-out instead of the predicted number  
3 of copies/ $\mu\text{L}$ , as these more accurately reflect the measurements made by the instrument and  
4 its technical biases. In this linear mixed-effects model, the FIR stage was the main explanatory  
5 variable, and aside from the random effect of the newborn, other covariates that entered as  
6 main factors in the model were sex, maximum lymphocyte count during the first postnatal week,  
7 and a batch indicator variable identifying the run of the TREC measurement. This linear mixed-  
8 effects model, as well as a null one without the FIR stage, were fitted to the data using the R  
9 package lme4<sup>68</sup>.

10  
11 A chi-squared test with one degree of freedom between the two models gave a  $P$  value=0.0432,  
12 thus rejecting the hypothesis of no FIR effect at a significance level  $\alpha < 0.05$ . Adjusted and  
13 corrected fluorescent counts on a logarithmic scale (shown in Figure 4a) were obtained after  
14 removing the effect of the previous covariates using a linear model.

## 15 **RNA extraction, library preparation and sequencing**

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

## 2 **Pre-processing and differential expression analysis of transcriptomics RNA-seq data**

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

10

11 Raw paired-end reads in FASTQ files, including the downsampled version of BS13, were  
12 aligned to the GRCh38 version of the reference human genome, without alternate locus  
13 scaffolds (GCA\_000001405.15) and including human decoy sequences from hs38d1  
14 (GCA\_000786075.2), using STAR<sup>69</sup> version 2.6.0c with default parameters except for --  
15 `peOverlapNbasesMin 10` and `--sjdbOverhang 74`, producing an initial set of 84 BAM  
16 files. Aligned reads in BAM files were reduced to a table of counts of 25,221 genes by 84  
17 samples using gene annotations from GENCODE v24 and the R/Bioconductor package  
18 *GenomeAlignments*<sup>70</sup> version 1.16.0, and its function `summarizeOverlaps`. We used specific  
19 arguments in the call to this function to restrict the count of genic reads to only those that fell  
20 entirely within the exonic regions and aligned to a unique site on the genome, to reflect library  
21 preparation protocols and avoid counting reads without a matching pair or overlapping multiple  
22 features.

23

24 The rest of the analysis of transcriptomics data was based on the *edgeR*<sup>71</sup> and *limma*<sup>72</sup>  
25 pipelines, and a few other R packages from the Bioconductor project<sup>73</sup> release version 3.7. A  
26 multidimensional scaling (MDS) plot of the samples showed a nearly perfect agreement  
27 between libraries from the same sample sequenced in each of the 4 different lanes. For this  
28 reason, we merged BAM files with aligned reads from the same sample and built a new table of  
29 counts of 25,221 genes by 21 samples, using the previously employed reduction procedure.  
30 Considering that 6 out of the 21 infants were not diagnosed with MIR or FIR (Supp. Fig. S1),  
31 and following previously established recommendations<sup>74</sup>, we filtered out lowly-expressed genes  
32 by discarding those that did not show a minimum reliable level of expression of 10 counts per  
33 million reads of the smallest library size, in less than 6 samples. This filtering step led to a final  
34 table of counts of 11,279 genes by 21 samples.

35

2 We used TMM normalization<sup>75</sup> to produce the MDS plot shown in Figure 1c, as well as to obtain  
3 gene-level robust dispersion estimates<sup>76</sup> that helped to select the sex of the infant and surrogate  
4 variables calculated with SVA<sup>77</sup>, as covariates in the linear models employed for differential  
5 expression analysis. Finally, we used the limma-voom<sup>78</sup> pipeline with quality weights, combined  
6 with quantile normalization, to conduct the differential expression analysis using linear models.  
7 The design matrix included the FIR status of the infants as a main explanatory variable and sex  
8 and surrogate variables, estimated with SVA, as covariates.

9  
10 The differential expression analysis across different combined stages of inflammation in MIR  
11 and FIR (Supp. Fig. S6, Supp. Table S4) was based on our replacing the FIR status variable in  
12 the linear model, with numerical values for MIR and FIR. Specifically, we assigned a 0 when  
13 MIR and FIR were not present, 1 when MIR was present but FIR was not, 2 when MIR was  
14 stage III and FIR was stage I and, finally, 3 when MIR was stage III and FIR was stage II or III.

#### 15 **Pre-processing and differential expression analysis of transcriptomics microarray data**

16 We obtained from Nigel Paneth and collaborators at the Michigan State University raw  
17 microarray gene expression data files from paired frozen and unfrozen DBS samples from nine  
18 neonates of the US ELGAN cohort, where two were affected by FIR and seven were not. The  
19 n=18 corresponding samples had been hybridized on a Agilent SurePrint G3 Human Gene  
20 Expression 8x60K Microarray Kit, and their analysis comparing the number of expressed genes  
21 between frozen and unfrozen samples was published by Wei *et al.* (2014)<sup>15</sup>. Batch processing  
22 information was derived from the scanning timestamp stored in raw data files as previously  
23 described<sup>79</sup>, creating a batch indicator variable that divided the samples into two different  
24 batches. The cross-classification of infants by FIR status and batch indicator showed no  
25 correlation between the primary outcome of our analysis, FIR status, and sample batch  
26 processing (Supp. Fig. S4). The MDS plot of the data shows the effect of storage temperature  
27 and batch processing, but not sex (Supp. Fig. S4).

28  
29 We corrected background expression intensities using the “normexp” method in the function  
30 backgroundCorrect from the limma package<sup>80</sup>, and used a quantile normalization approach to  
31 obtain comparable gene expression distributions across the samples. We removed probesets  
32 without annotation to Entrez gene identifiers, and probesets annotated to a common Entrez  
33 gene, keeping the one with highest variability measured by interquartile range (IQR). We then  
34 discarded the Agilent positive control probes and retained initially the negative control probes to



2 later estimate the fraction of expressed genes per sample. This resulted in an expression data  
3 set of 21,952 probesets in one-to-one correspondence with Entrez genes after discarding  
4 negative control probes, by  $n = 18$  samples. We calculated the 95% quantile of the expression  
5 level distribution observed in negative control probes for each sample. As a function of the  
6 percentage of increase over the 95% quantile of negative probe expression (from 100% to  
7 200% increasing by 10%), we found the number of expressed genes per sample separately by  
8 storage condition, which we later used for non-specific filtering based on minimum expression.

9  
10 Using the software package SVA<sup>77</sup> and the top 1% most variable (IQR) genes, we estimated  
11 one surrogate variable covariate that captures variability unrelated to FIR, storage temperature  
12 and batch processing. For each gene, we defined a linear mixed-effects model where the  
13 expression profile of every gene is a linear function of FIR status (yes/no, main explanatory  
14 variable), storage temperature (frozen/unfrozen), batch processing, the estimated surrogate  
15 variable covariate and the infant as a random effect. We estimated array quality weights to  
16 down-weight samples of worse quality in a graduated way<sup>81</sup> and to avoid discarding them. Using  
17 limma, we fitted this linear mixed-effects model, including the array quality weights, to each  
18 gene expression profile. We then we calculated moderated  $t$ -statistics, squeezing the genewise  
19 residual variances towards a global trend, for the coefficient estimating the effect between FIR  
20 and non-FIR-affected infants, and their corresponding  $P$  values for the null hypothesis of no-  
21 differential expression. We selected initially different subsets of DE genes at 5% FDR according  
22 to different thresholds of the non-specific filtering on minimum expression and, within each non-  
23 specific filter, FDR-adjusted  $P$  values were re-calculated. We found that a 110% cutoff yielded  
24 the highest number of DE genes, and using this non-specific filter we readjusted the raw  $P$   
25 values and obtained 33 genes that changed significantly at 5% FDR by 50% or more, between  
26 newborns affected and unaffected by FIR.

## 27 **Mass spectrometry proteomics**

28 *Sample preparation for mass spectrometry:* DBS samples were excised and collected in a 2 mL  
29 eppendorf, and extracted as described previously<sup>82</sup>. Briefly, each sample was soaked with 970  
30  $\mu\text{l}$  of a solution consisting of 758  $\mu\text{l}$  of 25 mM ammonium bicarbonate, 113  $\mu\text{l}$  of 10% (w/v)  
31 sodium deoxycholate and 99  $\mu\text{l}$  of 5 mM tris(2-carboxyethyl)phosphine, which was incubated at  
32 60 °C during 1h in constant agitation. Samples were alkylated with 10  $\mu\text{M}$  iodoacetamide (37°C,  
33 30 min) and quenched with 10  $\mu\text{M}$  dithiothreitol (37°C, 30 min). Samples were digested with  
34 trypsin (40 ng/ $\mu\text{l}$ ) overnight at 37 °C. Digestion was stopped with formic acid, and samples were

2 vortexed and centrifuged (800 xg, 20 min) prior to supernatant desalting using spin C18  
3 columns.

4  
5 *Mass spectrometry data acquisition:* Samples were reconstituted in 0.1% formic acid and were  
6 analyzed on a Orbitrap Fusion Lumos with an EASY-Spray nanosource coupled to a nano-  
7 UPLC system (EASY-nanoLC 1000 liquid chromatograph) equipped with a 50-cm C18 column  
8 (EASY-Spray; 75 µm id, PepMap RSLC C18, 2 mm particles, 45 °C). Chromatographic  
9 gradients were started at 5% buffer B with a flow rate of 300 nL/min and gradually increased to  
10 22% buffer B in 79 min and to 32% in 11 minutes. After each analysis, the column was washed  
11 for 10 min with 95% buffer B (Buffer A: 0.1% formic acid in water. Buffer B: 0.1% formic acid in  
12 acetonitrile). The mass spectrometer was operated in data-dependent acquisition mode, with full  
13 MS scans over a mass range of m/z 350–1500 with detection in the Orbitrap (120K resolution)  
14 and with auto gain control (AGC) set to 100,000. In each cycle of data-dependent acquisition  
15 analysis, following each survey scan, the most intense ions above a threshold ion count of  
16 10,000 were selected for fragmentation with HCD at normalized collision energy of 28%. The  
17 number of selected precursor ions for fragmentation was determined by the “Top Speed”  
18 acquisition algorithm (max cycle time of 3 seconds), and a dynamic exclusion of 60 s was set.  
19 Fragment ion spectra were acquired in the ion trap with an AGC of 10,000 and a maximum  
20 injection time of 200 ms. One of the samples, BS04, could not be profiled for technical reasons.  
21 All data were acquired with Xcalibur software v4.1.

22  
23 *Mass spectrometry protein expression quantification:* The MaxQuant software suite (v1.6.2.6)  
24 was used for peptide identification and label-free protein quantification<sup>83</sup>. The data were  
25 searched against the Uniprot human database with decoy entries (as of April 2018, 42,518  
26 entries). A precursor ion mass tolerance of 4.5 ppm at the MS1 level was used, and up to two  
27 missed cleavages for trypsin were allowed. The fragment ion mass tolerance was set to 0.5 Da.  
28 Oxidation of methionine, protein acetylation at the N-terminal were defined as variable  
29 modification; whereas carbamidomethylation on cysteines was set as a fixed modification. The  
30 minimum number of razor and unique peptides (“min. razor + unique”) for a protein group to be  
31 considered as identified was set to 1. Identified peptides and proteins were filtered using a 5%  
32 FDR. The match between runs algorithm was activated with a tolerance of 0.7 min for the  
33 matching time window and 20 min for the alignment time window. The mass spectrometry  
34 proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE<sup>84</sup>  
35 partner repository with accession number PXD011626.



## 2 **Pre-processing and differential expression analysis of proteomics data**

3 From the initial set of 649 quantified protein expression profiles on 20 samples, we discarded 26  
4 whose protein identifier could not be mapped to one of the 25,221 genes profiled by RNA-seq.  
5 From the remaining 624 proteins, we selected 303 with protein quantification values in at least 6  
6 out of the 20 samples. We further discarded 48 proteins whose protein quantification values  
7 were based on one single peptide, leading to a final protein quantification data matrix of 245  
8 proteins by 20 samples. Raw protein quantifications were normalized using the *vsn*<sup>85,86</sup>  
9 R/Bioconductor package. A complete normalized matrix of protein expression values was  
10 obtained by imputing missing quantifications with a multivariate method<sup>87</sup> based on the  
11 expectation-maximization (EM) algorithm implemented in the function *impute.wrapper.MLE* from  
12 the R package *imputeLCMD*. Using this complete and normalized matrix of protein expression  
13 values as an input to the *edgeR*<sup>71</sup> and *limma-trend*<sup>76</sup> pipelines, we produced the MDS plot in  
14 Figure 1d and conducted a protein differential expression analysis with linear models shown in  
15 Figure 2c. The design matrix included the FIR status of infants as the main explanatory variable  
16 and sex as a covariate.

## 17 **Functional enrichment analysis with the Gene Ontology database**

18 We performed functional enrichment analysis with the GO database using the R/Bioconductor  
19 package *GOstats*<sup>88</sup>. More concretely, we used the biological process ontology and the  
20 conditional hypergeometric test to account for dependencies derived from the GO hierarchical  
21 structure. Under this procedure, the GO hierarchy is traversed towards the root, applying a one-  
22 tailed Fisher's exact test between the genes in the GO term and the DE gene set at hand. When  
23 a GO term is considered to be significantly enriched by the DE gene set, its genes are removed  
24 from parent GO terms before testing them. To consider a GO term as significantly enriched we  
25 used a *P* value cutoff of 0.01, a minimum OR of 1.5, a minimum of five and a maximum of 300  
26 genes in the GO term. The gene universe was defined by the entire set of 25,221 profiled genes  
27 by RNA-seq. To report only the most robustly enriched biological processes, the final set of  
28 significantly enriched GO terms was selected by first using a multiple testing correction of FDR  
29 < 10% on the whole set of tested GO terms, then by further selecting GO terms with OR > 1.5  
30 and a minimum number of five genes enriching the GO term.

## 31 **Acknowledgments**

32 This work was supported by the Spanish MINECO/FEDER (TIN2015-71079-P). The CRG/UPF  
33 Proteomics Unit is part of the Spanish Infrastructure for Omics Technologies (ICTS Omics Tech)

2 and is a member of Proteored, PRB3, which is supported by grant PT17/0019 of the PE I+D+i  
3 2013-2016, funded by ISCIII and ERDF. We acknowledge support from the Spanish Ministry of  
4 Economy and Competitiveness, “Centro de Excelencia Severo Ochoa 2013-2017” (SEV-2012-  
5 0208), “Unidad de Excelencia Maria de Maeztu” (MDM-2014-0370), and “Secretaria  
6 d’Universitats i Recerca del Departament d’Economia i Coneixement de la Generalitat de  
7 Catalunya” (SGR17-595; SGR17-1020). The authors thank the ELGAN PAD committee, and  
8 Nigel Paneth, Sok Kean Khoo, Pete Haak, Madeleine Lenski and Elizabeth Allred, for granting  
9 access to the microarray expression and clinical data from the US ELGAN cohort. We would like  
10 to thank Aida Andrés, José Aramburu, Sergi Castellano, Alan Leviton and Cristina López-  
11 Rodríguez for their critical remarks on various parts of the manuscript.

## 12 **Contributors**

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

## 16 **Competing interests**

17 The authors declare no competing interests.

## 18 **Ethics approval**

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

## 22 **Data sharing statement**

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

## 27 **References**

28 1. Khong, T. Y. *et al.* Sampling and Definitions of Placental Lesions: Amsterdam Placental

- 2 Workshop Group Consensus Statement. *Arch. Pathol. Lab. Med.* **140**, 698–713 (2016).
- 3 2. Gantert, M. *et al.* Chorioamnionitis: a multiorgan disease of the fetus? *J. Perinatol.* **30**
- 4 **Suppl**, S21–30 (2010).
- 5 3. Dammann, O. & Leviton, A. Role of the fetus in perinatal infection and neonatal brain
- 6 damage. *Curr. Opin. Pediatr.* **12**, 99–104 (2000).
- 7 4. Costa, D. & Castelo, R. Umbilical cord gene expression reveals the molecular architecture
- 8 of the fetal inflammatory response in extremely preterm newborns. *Pediatr. Res.* **79**, 473–
- 9 481 (2016).
- 10 5. Tilley, S. K. *et al.* Genomic biomarkers of prenatal intrauterine inflammation in umbilical
- 11 cord tissue predict later life neurological outcomes. *PLoS One* **12**, e0176953 (2017).
- 12 6. Leviton, A. *et al.* Persistence after birth of systemic inflammation associated with umbilical
- 13 cord inflammation. *J. Reprod. Immunol.* **90**, 235–243 (2011).
- 14 7. O’Shea, T. M. *et al.* Inflammation-initiating illnesses, inflammation-related proteins, and
- 15 cognitive impairment in extremely preterm infants. *Brain Behav. Immun.* **29**, 104–112
- 16 (2013).
- 17 8. Dammann, O. & Leviton, A. Intermittent or sustained systemic inflammation and the
- 18 preterm brain. *Pediatr. Res.* **75**, 376–380 (2014).
- 19 9. Dammann, O. *et al.* Duration of Systemic Inflammation in the First Postnatal Month Among
- 20 Infants Born Before the 28th Week of Gestation. *Inflammation* **39**, 672–677 (2016).
- 21 10. Yanni, D. *et al.* Both antenatal and postnatal inflammation contribute information about the
- 22 risk of brain damage in extremely preterm newborns. *Pediatr. Res.* **82**, 691–696 (2017).
- 23 11. Korzeniewski, S. J. *et al.* A ‘multi-hit’ model of neonatal white matter injury: cumulative
- 24 contributions of chronic placental inflammation, acute fetal inflammation and postnatal
- 25 inflammatory events. *J. Perinat. Med.* **42**, 731–743 (2014).
- 26 12. Olin, A. *et al.* Stereotypic Immune System Development in Newborn Children. *Cell* **174**,
- 27 1277–1292.e14 (2018).

- 2 13. Papile, L. A., Burstein, J., Burstein, R. & Koffler, H. Incidence and evolution of  
3 subependymal and intraventricular hemorrhage: a study of infants with birth weights less  
4 than 1,500 gm. *J. Pediatr.* **92**, 529–534 (1978).
- 5 14. Sawyer, S., Krause, J., Guschanski, K., Savolainen, V. & Pääbo, S. Temporal patterns of  
6 nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One* **7**, e34131  
7 (2012).
- 8 15. Wei, C. *et al.* Comparison of frozen and unfrozen blood spots for gene expression studies.  
9 *J. Pediatr.* **164**, 189–191.e1 (2014).
- 10 16. Canna, S. W. *et al.* Life-threatening NLRC4-associated hyperinflammation successfully  
11 treated with IL-18 inhibition. *J. Allergy Clin. Immunol.* **139**, 1698–1701 (2017).
- 12 17. Condamine, T. *et al.* Lectin-type oxidized LDL receptor-1 distinguishes population of human  
13 polymorphonuclear myeloid-derived suppressor cells in cancer patients. *Sci Immunol* **1**,  
14 (2016).
- 15 18. Breuer, K. *et al.* InnateDB: systems biology of innate immunity and beyond—recent updates  
16 and continuing curation. *Nucleic Acids Res.* **41**, D1228–33 (2013).
- 17 19. Duncan, J. A. & Canna, S. W. The NLRC4 Inflammasome. *Immunol. Rev.* **281**, 115–123  
18 (2017).
- 19 20. Naranbhai, V. *et al.* Genomic modulators of gene expression in human neutrophils. *Nat.*  
20 *Commun.* **6**, 7545 (2015).
- 21 21. Boyman, O., Krieg, C., Homann, D. & Sprent, J. Homeostatic maintenance of T cells and  
22 natural killer cells. *Cell. Mol. Life Sci.* **69**, 1597–1608 (2012).
- 23 22. Haines, C. J. *et al.* Human CD4 T cell recent thymic emigrants are identified by protein  
24 tyrosine kinase 7 and have reduced immune function. *J. Exp. Med.* **206**, 275–285 (2009).
- 25 23. Lewis, D. B., Haines, C. & Ross, D. Protein tyrosine kinase 7: a novel surface marker for  
26 human recent thymic emigrants with potential clinical utility. *J. Perinatol.* **31 Suppl 1**, S72–  
27 81 (2011).

- 2 24. Pérès, M., Montfort, A., Andrieu-Abadie, N., Colacios, C. & Ségui, B. S1P: the elixir of life  
3 for naive T cells. *Cell. Mol. Immunol.* (2017). doi:10.1038/cmi.2017.110
- 4 25. Glavina-Durdov, M. *et al.* The grade of acute thymus involution in neonates correlates with  
5 the duration of acute illness and with the percentage of lymphocytes in peripheral blood  
6 smear. Pathological study. *Biol. Neonate* **83**, 229–234 (2003).
- 7 26. Schwarz, J. *et al.* Granulocytic myeloid-derived suppressor cells (GR-MDSC) accumulate in  
8 cord blood of preterm infants and remain elevated during the neonatal period. *Clin. Exp.*  
9 *Immunol.* **191**, 328–337 (2018).
- 10 27. Gervassi, A. *et al.* Myeloid derived suppressor cells are present at high frequency in  
11 neonates and suppress in vitro T cell responses. *PLoS One* **9**, e107816 (2014).
- 12 28. Leiber, A. *et al.* Neonatal myeloid derived suppressor cells show reduced apoptosis and  
13 immunosuppressive activity upon infection with Escherichia coli. *Eur. J. Immunol.* **47**,  
14 1009–1021 (2017).
- 15 29. Schieber, M. & Chandel, N. S. ROS function in redox signaling and oxidative stress. *Curr.*  
16 *Biol.* **24**, R453–62 (2014).
- 17 30. Veglia, F., Perego, M. & Gabrilovich, D. Myeloid-derived suppressor cells coming of age.  
18 *Nat. Immunol.* **19**, 108–119 (2018).
- 19 31. Zhou, J., Nefedova, Y., Lei, A. & Gabrilovich, D. Neutrophils and PMN-MDSC: Their  
20 biological role and interaction with stromal cells. *Semin. Immunol.* **35**, 19–28 (2018).
- 21 32. Ost, M. *et al.* Myeloid-Derived Suppressor Cells in Bacterial Infections. *Front. Cell. Infect.*  
22 *Microbiol.* **6**, (2016).
- 23 33. Tamadaho, R. S. E., Hoerauf, A. & Layland, L. E. Immunomodulatory effects of myeloid-  
24 derived suppressor cells in diseases: Role in cancer and infections. *Immunobiology* **223**,  
25 432–442 (2018).
- 26 34. Moskovitz, J. Roles of methionine sulfoxide reductases in antioxidant defense, protein  
27 regulation and survival. *Curr. Pharm. Des.* **11**, 1451–1457 (2005).

- 2 35. Wynn, J. L. *et al.* Targeting IL-17A attenuates neonatal sepsis mortality induced by IL-18.  
3 *Proc. Natl. Acad. Sci. U. S. A.* **113**, E2627–35 (2016).
- 4 36. Liang, J. *et al.* Novel NLRC4 Mutation Causes a Syndrome of Perinatal Autoinflammation  
5 With Hemophagocytic Lymphohistiocytosis, Hepatosplenomegaly, Fetal Thrombotic  
6 Vasculopathy, and Congenital Anemia and Ascites. *Pediatr. Dev. Pathol.* **20**, 498–505  
7 (2017).
- 8 37. Denes, A. *et al.* AIM2 and NLRC4 inflammasomes contribute with ASC to acute brain injury  
9 independently of NLRP3. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 4050–4055 (2015).
- 10 38. Ravichandran, S. *et al.* Higher daily doses of caffeine lowered the incidence of moderate to  
11 severe neurodevelopmental disabilities in very low birth weight infants. *Acta Paediatr.* **108**,  
12 430–435 (2019).
- 13 39. Furman, D. *et al.* Expression of specific inflammasome gene modules stratifies older  
14 individuals into two extreme clinical and immunological states. *Nat. Med.* **23**, 174–184  
15 (2017).
- 16 40. Hanna, N. *et al.* Mechanisms underlying reduced apoptosis in neonatal neutrophils. *Pediatr.*  
17 *Res.* **57**, 56–62 (2005).
- 18 41. Allgaier, B., Shi, M., Luo, D. & Koenig, J. M. Spontaneous and Fas-mediated apoptosis are  
19 diminished in umbilical cord blood neutrophils compared with adult neutrophils. *J. Leukoc.*  
20 *Biol.* **64**, 331–336 (1998).
- 21 42. Nguyen, C. N., Schnulle, P. M., Chegini, N., Luo, X. & Koenig, J. M. Neonatal neutrophils  
22 with prolonged survival secrete mediators associated with chronic inflammation.  
23 *Neonatology* **98**, 341–347 (2010).
- 24 43. Atallah, M. *et al.* Constitutive Neutrophil Apoptosis: Regulation by Cell Concentration via  
25 S100 A8/9 and the MEK – ERK Pathway. *PLoS One* **7**, e29333 (2012).
- 26 44. Hirata, J.-I. *et al.* A role for IL-18 in human neutrophil apoptosis. *Shock* **30**, 628–633 (2008).
- 27 45. Chen, K. W. *et al.* The neutrophil NLRC4 inflammasome selectively promotes IL-1 $\beta$

- 2 maturation without pyroptosis during acute Salmonella challenge. *Cell Rep.* **8**, 570–582  
3 (2014).
- 4 46. Kessel, C., Holzinger, D. & Foell, D. Phagocyte-derived S100 proteins in autoinflammation:  
5 putative role in pathogenesis and usefulness as biomarkers. *Clin. Immunol.* **147**, 229–241  
6 (2013).
- 7 47. Alder, M. N., Opoka, A. M., Lahni, P., Hildeman, D. A. & Wong, H. R. Olfactomedin-4 Is a  
8 Candidate Marker for a Pathogenic Neutrophil Subset in Septic Shock. *Crit. Care Med.* **45**,  
9 e426–e432 (2017).
- 10 48. Solan, P. D. *et al.* A novel role for matrix metalloproteinase-8 in sepsis. *Crit. Care Med.* **40**,  
11 379–387 (2012).
- 12 49. O’Hare, F. M. *et al.* Persistent systemic monocyte and neutrophil activation in neonatal  
13 encephalopathy. *J. Matern. Fetal. Neonatal Med.* **29**, 582–589 (2016).
- 14 50. Molloy, E. J. *et al.* Neonatal encephalopathy is associated with altered perinatal systemic  
15 neutrophil apoptosis. *Am. J. Perinatol.* **24**, 525–530 (2007).
- 16 51. Di Naro, E. *et al.* Fetal thymic involution: a sonographic marker of the fetal inflammatory  
17 response syndrome. *Am. J. Obstet. Gynecol.* **194**, 153–159 (2006).
- 18 52. Scheible, K. M. *et al.* T cell developmental arrest in former premature infants increases risk  
19 of respiratory morbidity later in infancy. *JCI Insight* **3**, (2018).
- 20 53. Kuban, J. D., Allred, E. N. & Leviton, A. Thymus involution and cerebral white matter  
21 damage in extremely low gestational age neonates. *Biol. Neonate* **90**, 252–257 (2006).
- 22 54. He, Y.-M. *et al.* Transitory presence of myeloid-derived suppressor cells in neonates is  
23 critical for control of inflammation. *Nat. Med.* **24**, 224–231 (2018).
- 24 55. Chemnitz, J. M. *et al.* Prostaglandin E2 impairs CD4+ T cell activation by inhibition of Ick:  
25 implications in Hodgkin’s lymphoma. *Cancer Res.* **66**, 1114–1122 (2006).
- 26 56. Ozsurekci, Y. & Aykac, K. Oxidative Stress Related Diseases in Newborns. *Oxid. Med.*  
27 *Cell. Longev.* **2016**, 1–9 (2016).



- 2 57. Tindell, R. & Tipple, T. Selenium: implications for outcomes in extremely preterm infants. *J.*  
3 *Perinatol.* **38**, 197–202 (2018).
- 4 58. Fong, O. N. *et al.* Expression profile of cord blood neutrophils and dysregulation of  
5 HSPA1A and OLR1 upon challenge by bacterial peptidoglycan. *J. Leukoc. Biol.* **95**, 169–  
6 178 (2014).
- 7 59. Kruzel, M. L., Zimecki, M. & Actor, J. K. Lactoferrin in a Context of Inflammation-Induced  
8 Pathology. *Front. Immunol.* **8**, (2017).
- 9 60. Song, J. & Kim, O. Y. Perspectives in Lipocalin-2: Emerging Biomarker for Medical  
10 Diagnosis and Prognosis for Alzheimer’s Disease. *Clin. Nutr. Res.* **7**, 1–10 (2018).
- 11 61. Yellowhair, T. R. *et al.* Preclinical chorioamnionitis dysregulates CXCL1/CXCR2 signaling  
12 throughout the placental-fetal-brain axis. *Exp. Neurol.* **301**, 110–119 (2018).
- 13 62. Denstaedt, S. J. *et al.* S100A8/A9 Drives Neuroinflammatory Priming and Protects against  
14 Anxiety-like Behavior after Sepsis. *J. Immunol.* **200**, 3188–3200 (2018).
- 15 63. Silk, E., Zhao, H., Weng, H. & Ma, D. The role of extracellular histone in organ injury. *Cell*  
16 *Death Dis.* **8**, e2812 (2017).
- 17 64. Gilthorpe, J. D. *et al.* Extracellular histone H1 is neurotoxic and drives a pro-inflammatory  
18 response in microglia. *F1000Res.* **2**, 148 (2013).
- 19 65. Hagberg, H. *et al.* The role of inflammation in perinatal brain injury. *Nat. Rev. Neurol.* **11**,  
20 192–208 (2015).
- 21 66. Gibbs, R. S. & Duff, P. Progress in pathogenesis and management of clinical intraamniotic  
22 infection. *Am. J. Obstet. Gynecol.* **164**, 1317–1326 (1991).
- 23 67. Redline, R. W. *et al.* Amniotic infection syndrome: nosology and reproducibility of placental  
24 reaction patterns. *Pediatr. Dev. Pathol.* **6**, 435–448 (2003).
- 25 68. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using  
26 lme4. *Journal of Statistical Software, Articles* **67**, 1–48 (2015).
- 27 69. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).



- 2 70. Lawrence, M. *et al.* Software for computing and annotating genomic ranges. *PLoS Comput.*  
3 *Biol.* **9**, e1003118 (2013).
- 4 71. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for  
5 differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140  
6 (2010).
- 7 72. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and  
8 microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
- 9 73. Huber, W. *et al.* Orchestrating high-throughput genomic analysis with Bioconductor. *Nat.*  
10 *Methods* **12**, 115–121 (2015).
- 11 74. Chen, Y., Lun, A. T. L. & Smyth, G. K. From reads to genes to pathways: differential  
12 expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-  
13 likelihood pipeline. *F1000Res.* **5**, 1438 (2016).
- 14 75. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression  
15 analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
- 16 76. Phipson, B., Lee, S., Majewski, I. J., Alexander, W. S. & Smyth, G. K. Robust  
17 hyperparameter estimation protects against hypervariable genes and improves power to  
18 detect differential expression. *Ann. Appl. Stat.* **10**, 946–963 (2016).
- 19 77. Leek, J. T. & Storey, J. D. Capturing heterogeneity in gene expression studies by surrogate  
20 variable analysis. *PLoS Genet.* **3**, 1724–1735 (2007).
- 21 78. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model  
22 analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
- 23 79. Leek, J. T. *et al.* Tackling the widespread and critical impact of batch effects in high-  
24 throughput data. *Nat. Rev. Genet.* **11**, 733–739 (2010).
- 25 80. Silver, J. D., Ritchie, M. E. & Smyth, G. K. Microarray background correction: maximum  
26 likelihood estimation for the normal-exponential convolution. *Biostatistics* **10**, 352–363  
27 (2009).

- 2 81. Ritchie, M. E. *et al.* Empirical array quality weights in the analysis of microarray data. *BMC*  
3 *Bioinformatics* **7**, 261 (2006).
- 4 82. Chambers, A. G., Percy, A. J., Yang, J. & Borchers, C. H. Multiple Reaction Monitoring  
5 Enables Precise Quantification of 97 Proteins in Dried Blood Spots. *Mol. Cell. Proteomics*  
6 **14**, 3094–3104 (2015).
- 7 83. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized  
8 p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.*  
9 **26**, 1367–1372 (2008).
- 10 84. Vizcaíno, J. A. *et al.* The PRoteomics IDentifications (PRIDE) database and associated  
11 tools: status in 2013. *Nucleic Acids Res.* **41**, D1063–9 (2013).
- 12 85. Huber, W., von Heydebreck, A., Sülthmann, H., Poustka, A. & Vingron, M. Variance  
13 stabilization applied to microarray data calibration and to the quantification of differential  
14 expression. *Bioinformatics* **18 Suppl 1**, S96–104 (2002).
- 15 86. Välikangas, T., Suomi, T. & Elo, L. L. A systematic evaluation of normalization methods in  
16 quantitative label-free proteomics. *Brief. Bioinform.* **19**, 1–11 (2018).
- 17 87. Schafer, J. L. *Analysis of Incomplete Multivariate Data*. (CRC Press, 1997).
- 18 88. Falcon, S. & Gentleman, R. Using GOstats to test gene lists for GO term association.  
19 *Bioinformatics* **23**, 257–258 (2007).

20