Small volumes, deep insights: longitudinal plasma EV multi-omics in very preterm infants

Nicole Grinovero[1†], Francesca Antonini[1†], Martina Bartolucci[1], Lucilla Rossi[1], Gianvittorio Luria[1], Maurizio Bruschi[2,3], Sonia Spinelli[2], Gino Tripodi[4], Chiara Andreato[5,6], Francesco Vinci[5,6], Luca Antonio Ramenghi[5,6‡], Chiara Lavarello[1*‡], Andrea Petretto[1‡]

[1] Core Facility for Omics Sciences, IRCCS Istituto Giannina Gaslini, Genoa, Italy

[2] Unit of Nephrology, Dialysis, and Transplantation and Laboratory of Molecular Nephrology, IRCCS Istituto Giannina Gaslini, Genoa, Italy

[3] Department of Experimental Medicine (DIMES), University of Genoa, Italy

[4] Immunohematology and Transfusion Medicine Unit, IRCCS Istituto Giannina Gaslini, Genoa, Italy

[5] Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, University of Genoa, Italy

[6] Neonatal Intensive Care Unit, IRCCS Istituto Giannina Gaslini, Genoa, Italy

[†] These authors contributed equally to this work.

[‡] These authors jointly supervised this work.

* Correspondence to: Chiara Lavarello (chiaralavarello@gaslini.org)

Abstract

Very preterm birth disrupts critical fetal developmental programs, yet the systemic molecular trajectories driving extrauterine adaptation remain poorly defined. Although extracellular vesicles (EVs) represent informative systemic compartments, comprehensive multi-omics is constrained by the small plasma volumes safely obtainable from neonates. Here, we adapted a magnetic bead-based framework (Mag-Net) to enable parallel EV proteomics and lipidomics from the same EV-enriched preparation using 10 µL of plasma. Across 74 longitudinal samples collected from birth to term-equivalent age, we quantified 1,528 EV-associated proteins and 421 lipid species. The EV proteome shifted from early translation and metabolic programs toward progressive immune competence, while the lipidome underwent selective structural remodeling enriched in triacylglycerols and ether-linked phosphatidylcholines. Cross-omics integration identified coordinated protein-lipid modules associated with clinical phenotypes, including brain injury. This study demonstrates that parallel EV proteomic-lipidomic profiling from microliter plasma volumes is feasible and captures coordinated developmental and clinically relevant programs in very preterm infants.

Keywords: *Extracellular vesicles, Multi-omics, Preterm birth, Liquid biopsy, Brain injury*

Introduction

Preterm birth affects approximately 13-15 million infants each year worldwide and remains a leading cause of neonatal mortality and long-term neurodevelopmental impairment[1]. Infants born very preterm (<32 weeks of gestation) undergo extrauterine adaptation during the third trimester, a critical window marked by the rapid and coordinated maturation of neural, immune and metabolic systems[2,3]. This transition interrupts tightly regulated fetal programs, forcing maturation to unfold under fundamentally different biological conditions shaped by altered oxygen exposure, inflammation, and modified nutritional inputs. As a result, postnatal development in very preterm infants does not simply represent an accelerated version of normal maturation, but unfolds under different biological constraints[4,5].

During the first weeks of life, intrinsic maturation occurs alongside clinical exposures such as respiratory support, infection and hemodynamic instability. Because developmental timing and stress responses are tightly intertwined, cross-sectional comparisons provide limited insight into the molecular organization of early adaptation. In addition, inter-individual variability is substantial in very preterm cohorts, reflecting differences in gestational age and clinical course. Longitudinal designs, in which each infant serves as their own reference, are therefore essential to distinguish structured developmental trajectories from reactive changes linked to clinical events[6].

Extracellular vesicles (EVs) provide a biologically relevant analytically tractable compartment for this purpose. EVs are membrane-bound particles released by most cell types and mediate intercellular communication through the transfer of proteins, lipids and nucleic acids[7,8]. As circulating, cell-derived structures, they integrate signals across tissues and can be accessed through minimally invasive plasma sampling, making them well suited for longitudinal studies in neonatal populations[9,10]. Importantly, EVs are structured molecular entities. Their lipid bilayer is selectively assembled and enriched in specific phospholipids and sphingolipids that influence vesicle curvature, stability and fusion, while membrane-associated and luminal proteins regulate trafficking and signaling[11]. Because membrane lipids and proteins are functionally linked, parallel analysis of EV proteomes and lipidomes enables an integrated view of both structural and signaling aspects of systemic adaptation.

Despite this potential, EV analysis in plasma poses well-recognized technical challenges. EVs represent only a small fraction of circulating particles and are outnumbered by lipoproteins and abundant soluble proteins. Overlap in size and density complicates separation, and incomplete resolution is particularly problematic for lipidomics, where co-isolated lipoprotein lipids can dominate the measured signal[12,13]. Benchmarking studies have shown that isolation strategies

involve trade-offs between purity and yield: precipitation-based methods often provide higher recovery but increased contamination, whereas density-gradient or size-exclusion approaches improve purity but reduce yield and require larger input volumes[12].

These limitations are amplified in neonatal research, where available plasma volumes are inherently limited and repeated sampling must remain minimally invasive to allow longitudinal follow-up. Many EV isolation workflows require hundreds of microliters to milliliters of plasma and multiple processing steps, making them poorly suited for fragile populations. In addition, most EV multi-omics studies process proteomics and lipidomics separately, dividing enriched material into aliquots. When starting volumes are small, this reduces analytical depth and disrupts the natural coupling between membrane composition and protein cargo. Although parallel EV proteomics and lipidomics have been demonstrated in adult plasma using larger-volume, density-gradient-based workflows[13], integrated proteomic-lipidomic profiling from microliter-scale plasma inputs in a longitudinal neonatal setting has not been established.

To address part of the isolation challenge, a strong anion exchange (SAX)-based enrichment strategy (Mag-Net) was recently developed for plasma proteomics[14]. By exploiting the net negative surface charge of EV membranes, Mag-Net enables selective enrichment of vesicle-associated material from complex plasma matrices while reducing interference from highly abundant soluble proteins. This workflow has been validated for deep plasma EV proteome profiling and provides an automated, scalable platform compatible with clinical samples[14].

In this study, we build on the Mag-Net framework and extend it to enable integrated proteomic and lipidomic analysis from plasma micro-volumes. We evaluate and refine the workflow using 5 and 10 µL plasma inputs, volumes compatible with longitudinal sampling in very preterm infants, and optimize the downstream extraction strategy to enable monophasic recovery of proteins and lipids from the same EV preparation. This approach preserves cross-layer molecular relationships without splitting the input material and demonstrates the feasibility of parallel EV multi-omics from microliter-scale plasma volumes.

After technical validation and benchmarking against conserved circulating EV signatures[13], we apply this optimized workflow longitudinally in a clinically characterized cohort of very preterm infants from birth to term-equivalent age. Using linear mixed-effects modeling and cross-omics network integration, we characterize coordinated EV molecular trajectories during early extrauterine adaptation and examine their associations with clinical outcomes, with particular attention to brain injury.

Results

Optimization of an integrated EV proteomic and lipidomic workflow from plasma micro-volumes

Integrated clinical multi-omics in fragile populations, such as neonates or patients with rare diseases, is often constrained by minimal sample availability. To enable simultaneous protein and lipid profiling under these micro-volume conditions, we adapted the previously established Mag-Net workflow[14] – originally validated for plasma proteomics – to support parallel analysis from as little as 5-10 µL of plasma. While preserving the upstream strong anion exchange (SAX)-based EV enrichment step, optimization focused on the downstream extraction strategy to ensure compatibility with both analytes. We compared three monophasic extraction chemistries: methanol (MeOH), isopropanol (IPA), and a 1-butanol/acetonitrile/water mixture (3:1:1, v/v), based on the Bead-enabled Accelerated Monophasic Multi-omics (BAMM) approach[15]. In all conditions, solvent-mediated membrane disruption was coupled with on-bead protein aggregation capture (PAC), allowing coordinated recovery of lipids and proteins from the same EV-enriched preparation.

Proteomic depth remained stable across extraction strategies and input volumes (**Fig. 1a**). At 10 µL of plasma input, we identified 866-992 protein groups, with only minor losses in coverage when scaling down to 5 µL. In contrast, direct digestion of whole plasma ("neat") yielded markedly lower depth (308 protein groups), confirming that SAX-based enrichment significantly increases the detectable complexity of the EV-associated proteome. Quantitative reproducibility was consistently high, with most proteins exhibiting a coefficient of variation (CV) < 20% (**Fig. 1b**). Furthermore, concordance with the standard Mag-Net implementation remained strong (Pearson r = 0.89-0.95; **Fig. 1c**), and the preservation of the global EV-associated proteomic landscape was further supported by correlation structure and protein overlap analysis (**Fig. S1**).

Conversely, lipidomic performance was more sensitive to the choice of extraction solvent. Across conditions, we detected up to 308 lipid species in positive mode and 108 in negative mode (**Fig. S2**). While total feature counts were broadly comparable, restricting the analysis to quantitatively robust signals (intensity > $1\times10^{\square}$) revealed solvent-dependent differences (**Fig. 1e**; **Fig. S2**). BAMM extraction consistently yielded the highest lipid coverage in positive mode and the lowest CV values (**Fig. 1f**), reflecting superior quantitative reproducibility. Notably, lipid class distributions remained stable across solvents (**Fig. S2**), indicating that extraction chemistry primarily influenced detection depth rather than inducing systematic compositional bias.

To verify that these molecular profiles reflect selective EV enrichment rather than plasma carryover, we benchmarked our dataset against a conserved circulating EV atlas defined by Rai et al. using density-gradient fractionation[13]. Of the 182 highly conserved EV proteins reported in that study, 168 (92%) were identified in our dataset. These markers showed coherent enrichment in

Mag-Net preparations relative to direct plasma digestion, while abundant non-EV plasma proteins were strongly depleted (**Fig. 1d**). At the lipid level, 9 of the 52 conserved EV lipid features described in the reference atlas were detected; although the absolute overlap was lower than for proteins, these features were consistently detected across all Mag-Net conditions, with BAMM 10 µL displaying the most robust profile (**Fig. S2**). Collectively, these data demonstrate that the adapted Mag-Net-based workflow recapitulates a conserved EV molecular signature, with BAMM extraction at 10 µL input providing the optimal balance of lipidomic robustness and proteomic integrity.

Longitudinal evolution of the plasma EV proteome and lipidome across very preterm development

With the optimized EV multi-omics workflow established, we investigated how circulating EV molecular programs evolve during early extrauterine adaptation. We applied the method to a longitudinal cohort of 16 very preterm infants (<32 weeks gestational age) followed from birth to term-equivalent age (TEA; 40 weeks postmenstrual age). TEA serves as a critical clinical reference point, representing the stage at which a preterm infant reaches the developmental maturity of a term-born newborn.

Plasma samples were collected at five predefined timepoints: birth (T0), 48-72 hours (T1), 7 days (T2), 33 weeks postmenstrual age (T3), and TEA (T4). While all 16 infants contributed samples for the T0-T3 interval, 10 were followed up to the final TEA milestone. The longitudinal design and analytical workflow are summarized in **Fig. 2a**, and clinical characteristics are reported in **Table 1**. Parallel analysis of proteins and lipids from each sample enabled within-subject tracking of coordinated multi-omic trajectories over time. Based on the superior coverage and quantitative robustness observed during our technical benchmarking, longitudinal lipidomic profiling was prioritized in positive ionization mode to ensure maximum analytical depth across the cohort.

After quality filtering, we retained 1,528 EV-associated protein groups and 421 unique lipid species for longitudinal analysis. Post-filter quantification depth per sample remained stable across timepoints in both omics layers (**Fig. S3**), supporting consistent analytical coverage throughout the longitudinal series. The EV-associated proteome spanned a wide dynamic range (**Fig. 2b**) and showed significant over-representation of extracellular and vesicle-related compartments, consistent with its EV origin (**Fig. 2c**). In parallel, the lipidomic profile revealed a membrane-enriched composition dominated by glycerophospholipids and sphingolipids, with sphingomyelins as the predominant lipid class (**Fig. 2d**).

To characterize the systematic remodeling of proteomic and lipidomic profiles, we employed a linear mixed-effects modeling (LMM) framework. This approach accounted for the longitudinal design by incorporating infant-specific random intercepts to model within-subject correlations over time. Sampling timepoints (T0-T4) were treated as fixed effects to capture developmental trends,

while gestational age at birth was included as a covariate, as it represents a major source of biological heterogeneity in very preterm cohorts and strongly influences baseline molecular profiles. For each molecular feature, we tested for a global effect of time and quantified the primary developmental shift by contrasting TEA (T4) against birth (T0), using the Benjamini-Hochberg (BH) procedure to control for multiple testing.

Extensive postnatal remodeling of the circulating EV proteome

Applying the LMM framework revealed a profound temporal transformation of the plasma EV proteome following very preterm birth. Overall, 57% of the quantified proteins (875 of 1,528) exhibited a significant global effect of time (FDR < 0.05), indicating that the majority of EV cargo is actively and dynamically regulated during the first weeks of extrauterine life.

Unsupervised hierarchical clustering of these time-regulated proteins resolved four distinct clusters that follow two primary, opposing trajectories: proteins that progressively decreased from birth toward maturation (clusters 1A and 1B) and proteins that steadily increased over time (clusters 2A and 2B) (**Fig. 3a,b**). The presence of these highly coordinated trajectories suggests that longitudinal EV proteomic shifts are coherent and directional, rather than reflecting unspecific physiological variability.

To quantify the magnitude of this transition, we focused on the net developmental shift between birth and TEA. We identified 439 proteins significantly downregulated and 127 proteins significantly upregulated at TEA (FDR < 0.05 and $|\beta| > 0.58$, corresponding to approximately ≥1.5-fold change; **Fig. 3c**). Proteins that decreased toward TEA were central to RNA metabolism, translation, and early growth-related processes (**Fig. 3e**). This "early-life" signature included core components of the translational machinery, such as RPS6 and EIF4A2, alongside the classic fetal marker alpha-fetoprotein (AFP), all of which showed steep declines across the five timepoints (**Fig. 3d**).

In striking contrast, the proteomic profile at TEA was characterized by a mature systemic signature. Proteins increasing toward TEA were significantly enriched for immune-related functions, specifically the complement cascade, immunoglobulin-mediated responses, and adaptive immunity (**Fig. 3e**). Representative markers of this maturation – including complement component C9, the anticoagulant protein S (PROS1), and the lipid transporter MTTP – exhibited linear increases across postnatal development (**Fig. 3d**). Collectively, these findings highlight a coordinated functional transition of the EV proteome from a translation-heavy growth program at birth toward a robust immune-competent profile at term-equivalent age.

Selective structural and metabolic remodeling of the EV lipidome

We next evaluated the longitudinal evolution of the plasma EV lipidome using the same LMM framework. While lipidomic remodeling involved a more targeted subset of features compared to the proteome – with 100 of 421 quantified species showing a significant global effect of time (FDR

< 0.05) – the changes were highly structured and coordinated (**Fig. 4a**). Unsupervised hierarchical clustering resolved these regulated lipids into four distinct developmental trajectories, each comprising species from multiple chemical classes (**Fig. 4b**). This modular organization indicates that EV lipidomic shifts are governed by coordinated metabolic programs rather than uniform, class-wide abundance changes.

The developmental transition to TEA was marked by the significant upregulation of 41 lipid species and the downregulation of 22 (FDR < 0.05 and |β| > 0.58, corresponding to approximately ≥1.5-fold change; **Fig. 4c**). Lipids increasing over development were dominated by triacylglycerols (TG), ether-linked phosphatidylcholines (PC O-), and specific sphingomyelins (SM). Among the strongest increases were SM 28:1;O2 and ether-linked species such as PC O-39:8 and PC O-37:8, alongside multiple TG species (**Fig. 4c**). These shifts are consistent with a progressive enrichment of neutral lipid cargo and a selective remodeling of EV membrane architecture.

Conversely, lipids that declined toward TEA were significantly enriched for diacyl PCs, phosphatidylserines (PS), diacylglycerols (DG), and long-chain SMs. Representative examples included PC 32:1 and PC 31:0, which showed the strongest negative effect sizes, as well as long-chain SM such as SM 36:0;O2 and SM 44:1;O2. Notably, lipid regulation was highly structural-specific rather than class-uniform. Specifically, SMs displayed opposite trends depending on chain length, with shorter species increasing and longer species decreasing over development, while PCs had divergent behavior according to linkage type, with ether-linked species preferentially increasing and diacyl species preferentially decreasing toward TEA.

To place these compositional changes into a functional context, we performed lipid ontology (LION) enrichment analysis based on the TEA vs birth β coefficients ($\log_2$ fold-change estimates) (**Fig. 4d**). Lipids enriched at T4 were associated with annotations related to lipid storage and lipid droplet biology. In contrast, lipids reduced over time were preferentially associated with membrane structural properties, including bilayer thickness and intrinsic curvature.

Finally, we mapped the regulated species from major membrane classes (PC, PC O-, and SM) onto carbon number-double bond space (**Fig. 4e**). This analysis revealed that birth- and TEA-enriched species occupy distinct regions of structural space within each class. Importantly, these shifts were class-specific and did not reflect a uniform global trend toward increased chain length or saturation, but rather a selective redistribution among molecular species.

Together, these analyses demonstrate that postnatal development in very preterm infants is characterized by a progressive and highly coordinated remodeling of the plasma EV composition across both omics layers. While proteomic changes are widespread and involve fundamental shifts in pathways related to growth and immune maturation, lipidomic changes are more selective, reflecting a sophisticated redistribution among specific structural species. Overall, these findings establish that plasma EVs serve as high-resolution molecular snapshots of the systemic

developmental changes occurring during the critical preterm period, providing a robust platform for tracking extrauterine adaptation.

Network-based integration reveals coordinated multi-omic EV modules

To characterize the systemic architecture of EV maturation, we integrated the proteomic and lipidomic layers using a cross-omics correlation network approach (**Fig. 5a**). We focused on the dynamically regulated components of EV remodeling by including only the 875 proteins and 100 lipids that exhibited significant temporal trajectories in our longitudinal LMM analysis. Inter-omic coordination was assessed by computing pairwise Spearman correlations (ρ) across paired longitudinal samples. By restricting the network to cross-layer associations and to features with a significant time effect (FDR < 0.05), we intentionally captured coordinated protein–lipid co-variation along the shared developmental trajectory, rather than within-layer structure or correlations driven by absolute abundance. After multiple-testing correction and effect-size filtering ($|\rho| \geq 0.35$, FDR < 0.05), the resulting network revealed a bimodal distribution of associations, with clear topological separation between positive and negative correlations (**Fig. 5b**).

Using Louvain community detection, we resolved the network into five distinct cross-omic modules (M1-M5), each representing a set of proteins and lipids with coherent longitudinal behavior (**Fig. 5a, c**). While proteins constituted the majority of nodes, lipids were consistently integrated within each module rather than forming isolated clusters, highlighting the tight biochemical coupling of the EV multi-ome. Functional annotation of protein components within each module identified distinct biological hallmarks of preterm development (**Fig. 5d**). Module M1, the largest community, reflected a systemic shift in neonatal metabolism through its enrichment in macromolecule catabolism and intracellular organization, effectively capturing the progressive decline of growth-associated cargo as the infant adapts to the extrauterine environment. The maturation of the neonatal immune system was divided into two distinct programs: module M2 was characterized by endocytosis and complement-mediated responses, hallmarks of the innate immune system, while module M3 focused on the expansion of the adaptive immune repertoire, specifically immunoglobulin-mediated signaling. The remaining modules highlighted the biosynthetic and structural demands of development; module M4 was associated with ribosomal biogenesis and cytoplasmic translation, whereas module M5 centered on extracellular matrix organization and tissue remodeling, suggesting a direct link between EV molecular cargo and the maturation of systemic tissue architecture.

Association of integrated EV molecular modules with clinical phenotypes

To evaluate the clinical relevance of the integrated EV programs, we interrogated the associations between cross-omics molecular modules and neonatal outcomes using a CAMERA-based

competitive gene set framework (**Fig. 6a**). By incorporating linear mixed-effects modeling with duplicate correlation, we accounted for the longitudinal nature of the data while adjusting for gestational age at birth. This module-level strategy increased statistical power to detect coordinated molecular programs that remained below the significance threshold in single-feature analyses.

Module M1 displayed a highly significant association with hemorrhagic brain lesions (FDR = $1.6 \times 10^{-}$□). Although the global module trajectory reflected a progressive decline in metabolic and proteostatic cargo, its leading-edge signal was driven by a compensatory increase in complement regulators C4BPA (β = 1.44) and C4BPB (β = 0.82), contrasted with a marked depletion of the hemoglobin subunit HBM (β = −0.74) (**Fig. 6b**). This internal divergence suggests selective remodeling of EV cargo in the context of hemorrhagic injury, with complement modulation emerging as a dominant adaptive feature.

In parallel, Module M3 exhibited consistent positive associations with both hemorrhagic lesions and the broader category of any brain lesion, including white matter injury (FDR = $3.7 \times 10^{-}$□) (**Fig. 6a,c**). This module reflected a sustained neuroinflammatory and adaptive immune signature, characterized by coordinated upregulation of immunoglobulin families and stress-response mediators such as NIBAN1 (β = 0.55) and the immunoproteasome subunit PSMB8 (β = 0.50), alongside reduced levels of the adhesion molecule SELL (β = −0.49). These trajectories indicate altered leukocyte activation and trafficking dynamics associated with brain injury.

The module-level framework further uncovered concerted lipidomic shifts that were not detectable at the single-feature level (**Fig. 6a,d**). Module M2 lipids were negatively associated with hemorrhagic lesions (FDR = 0.02), reflecting coordinated reductions in membrane-related species, including ether-linked phosphatidylcholines such as PC O-37:8 (β = −0.96) and PC O-39:8 (β = −0.79). Conversely, lipid components within module M3 were positively associated with systemic stress phenotypes, including sepsis (FDR = 0.04) and mechanical ventilation (FDR = 0.01). Representative species such as PC 33:1 and PC 32:1 showed sustained elevations in affected infants (**Fig. 6d**), accompanied by increases in diacylglycerols, consistent with lipid mobilization during inflammatory and respiratory stress.

Collectively, these results demonstrate that clinical phenotypes in very preterm infants are more robustly captured at the level of integrated EV molecular modules than individual features. Notably, the cross-layer convergence of Module M3 across proteins and lipids reinforces a coordinated link between adaptive immune activation and systemic inflammatory outcomes, highlighting the value of multi-omics integration to resolve distributed biological programs underlying neonatal pathology.

Discussion

Very preterm birth displaces a critical phase of third-trimester maturation into an extrauterine environment characterized by abrupt shifts in oxygen exposure, inflammation, nutrition and intensive care interventions. Longitudinal systems studies have shown that postnatal immune development follows highly ordered and reproducible trajectories, even in preterm infants[6,16,17]. Dissecting intrinsic developmental timing from superimposed clinical stress therefore requires molecular readouts that are both temporally sensitive and biologically specific. Here, we demonstrate that plasma EVs provide such a compartment. By extending Mag-Net enrichment[14] to enable integrated proteomic and lipidomic profiling from 10 µL plasma, and applying this framework from birth to term-equivalent age, we reconstruct structured EV trajectories and identify cross-omics modules associated with clinically relevant phenotypes, most prominently brain injury.

A first advance is methodological, with direct implications for biological interpretation. Plasma EV profiling is constrained not only by minimal input volumes but also by incomplete separation from abundant non-EV particles, particularly lipoproteins. High-resolution plasma fractionation studies now define conserved EV protein and lipid features alongside non-EV hallmarks[13], and demonstrate that lipoproteins can physically associate with EVs, forming hybrid nanoparticle complexes[18]. Methodological work further shows that combined size- and density-based approaches improve purity but cannot completely eliminate co-isolates[19]. Within this context, extracting proteins and lipids from the same EV preparation preserves the intrinsic coupling between membrane composition and protein cargo, strengthening cross-layer interpretation while acknowledging known plasma complexities.

Biologically, EV remodeling across early postnatal life is highly structured. More than half of quantified EV proteins display significant temporal dynamics, resolving into two dominant programs: a birth-enriched profile dominated by translation and RNA metabolism, and a progressively increasing immune program enriched for complement and immunoglobulin-associated functions. This transition mirrors systems-level neonatal studies demonstrating rapid, time-driven immune reorganization, with complement pathways among the most reproducibly regulated components in early life[6,20]. In large extremely preterm cohorts, postnatal age explains a substantial proportion of circulating proteomic variance, exceeding gestational age effects beyond the earliest window[17]. The EV trajectories observed here are therefore consistent with a broader temporal architecture of immune ontogeny, but captured within a selectively packaged vesicular compartment rather than bulk plasma.

The immune “ramp” we observe is not generic. Longitudinal plasma proteomics demonstrates coordinated early increases in classical and terminal complement components, rising IgM, and declining maternally derived IgG during the first week of life[20]. Our EV modules enriched for

complement and immunoglobulin-linked programs align with these dynamics, suggesting that EV cargo reflects structured maturation of innate and adaptive immune interfaces rather than nonspecific inflammation.

The EV lipidome follows a complementary but distinct pattern. Although fewer lipid species are time-structured, remodeling is clearly species- and structure-dependent. Enrichment toward TEA of triacylglycerols and ether-linked phosphatidylcholines, together with chain-length and linkage-specific redistribution within membrane classes, argues against uniform class-wide drift. Ether lipids are mechanistically linked to exosome biogenesis: modulation of ether-lipid synthesis alters vesicle release and composition[21], and plasmalogen enrichment influences membrane biophysical properties[22]. These data support the interpretation that ether-PC increases may reflect regulated membrane remodeling during EV maturation. At the same time, neutral lipid signals must be interpreted cautiously. LDL particles can co-purify and associate with EVs[18,23], and TG species are enriched in non-EV plasma fractions[13]. Conversely, lipid-rich exosome-sized vesicles with immuno-metabolic functions have been described[24]. Our interpretation therefore remains balanced: TG enrichment may reflect both regulated EV-associated biology and the evolving nanoparticle environment of neonatal plasma.

Integration of proteomic and lipidomic layers provides the clearest insight. By restricting network construction to temporally regulated features and modeling cross-layer correlations, we identify five modules summarizing coordinated longitudinal behavior. Network-based dimensionality reduction reduces multiple testing burden and yields biologically coherent units that are more stable than single features in small, heterogeneous cohorts. In our dataset, modules partition immune maturation into distinct programs and capture distributed lipid shifts that are not evident at the single-molecule level. This module-level organization transforms high-dimensional data into interpretable biological states.

Clinically, module-phenotype associations converge on brain injury. Complement provides a biologically plausible bridge between systemic inflammation and immature white matter vulnerability. In neonatal models, activation of the C3a-C3aR axis impairs oligodendrocyte precursor maturation and promotes hypomyelination[25]. Genetic disruption of classical pathway components such as C1q reduces hypoxic-ischemic brain injury in neonatal models[26,27], while C5a signaling contributes to cortical injury in experimental preterm birth[28]. Complement components also participate in glia-mediated neurotoxicity, including C1q-driven astrocyte activation[29]. These mechanistic data support the plausibility that EV modules enriched for complement components reflect systemic states linked to brain vulnerability.

Translational work further supports EV-mediated inter-organ communication in prematurity. In preterm infants, higher oxygen exposure associates with increased circulating EV cargo containing inflammasome components, and adoptive transfer of such EVs induces lung and brain injury

phenotypes in neonatal models[30]. EV cargo has also been linked to bronchopulmonary dysplasia severity with functional validation[31]. Together, these findings support the concept that circulating EVs can report and potentially mediate lung-immune-brain interactions. Within this framework, associations between our EV modules, respiratory support and brain injury are biologically coherent.

An additional interpretive layer concerns the plasma protein corona. EVs incubated in plasma acquire surface-associated proteins, including complement components and immunoglobulins[32]. This is particularly relevant in plasma, where EV surface-associated proteins may represent vesicle biology and the surrounding environment simultaneously. Importantly, such coronas modulate immune cell uptake and trafficking[33]. Complement- and Ig-rich modules may therefore reflect a combination of intrinsic cargo remodeling and altered opsonization dynamics, both relevant to neonatal inflammation and sepsis. Together, these observations indicate that immune-enriched EV modules provide a structured representation of vesicle-associated immune remodeling.

Several limitations define the scope of inference. Plasma EV preparations remain heterogeneous in subtype and cellular origin, and non-EV contributions cannot be fully excluded, particularly in lipid space. Our aim was to characterize an EV-enriched systemic compartment rather than assign tissue specificity, although EV proteomics has demonstrated tissue-enriched signatures in plasma under optimized conditions[34]. Nutritional transitions plausibly influence both immune and lipid trajectories. Although longitudinal mixed-effects modeling and module-level statistics mitigate limited cohort size and unbalanced late sampling, independent validation will be required before clinical translation.

In summary, integrated EV proteomic–lipidomic profiling from 10 µL of plasma captures coherent, time-structured molecular programs in very preterm infants. By enabling parallel recovery of proteins and lipids from the same micro-volume preparation, this workflow preserves cross-layer biological relationships under sampling conditions that are realistic in the NICU setting. EV proteins trace a coordinated shift from growth-associated translation toward immune competence, consistent with established patterns of neonatal immune ontogeny[6,16]. EV lipids undergo selective, structure-dependent remodeling compatible with regulated membrane biology[21], while remaining interpretable within the known constraints of plasma nanoparticle complexity. Cross-omics modules condense these longitudinal dynamics into stable and biologically meaningful units and reveal clinically relevant programs associated with brain injury and systemic stress. Together, these findings demonstrate that ultra-low-volume EV multi-omics is not only technically feasible, but capable of resolving structured developmental and injury-related trajectories, providing a scalable framework for longitudinal molecular profiling in neonatal medicine.

Methods

Cohort description

This prospective, longitudinal study included 16 very preterm infants born at <32 weeks of gestational age and admitted to the Neonatal Pathology and Intensive Care Unit of IRCCS Istituto Giannina Gaslini (Genoa, Italy). Infants were consecutively enrolled between November 2022 and November 2023 and followed from birth to term-equivalent age (40 weeks postmenstrual age) according to a predefined study protocol.

A total of 74 plasma samples were collected across five scheduled timepoints (T0-T4). Only infants with samples available at a minimum of four timepoints were included in longitudinal analyses.

Demographic and clinical data were collected from electronic medical records. Baseline characteristics were selected a priori based on relevance to neonatal maturation and availability in the clinical dataset. A summary of demographic and clinical characteristics is provided in **Table 1**.

Neurological status was assessed using routine clinical neuroimaging, including cranial ultrasound and, when clinically indicated, magnetic resonance imaging. Brain injury was defined as the presence of any documented brain lesion and recorded as a binary variable. When present, lesions were classified descriptively as hemorrhagic or white matter injury.

The study was conducted in accordance with the Declaration of Helsinki and approved by the Liguria Regional Ethics Committee (approval ID: 469/2021 – DB ID 11730). Written informed consent was obtained from the parents or legal guardians prior to enrollment.

Patients characteristics	
Gestational age at birth (weeks)	29.5 [28.5–31.2]
Birth weight (g)	1,255 [849–1,561]
Male sex	9 (56.2%)
Any brain lesion	8 (50.0%)
Hemorrhagic brain lesion	6 (37.5%)
White matter lesion	2 (12.5%)
Sepsis	8 (50.0%)
Mechanical ventilation	9 (56.2%)
Data are presented as median [Q1–Q3] or n (%). N = 16.	

Table 1: Baseline demographic and clinical characteristics of the study cohort (N = 16). Continuous variables are reported as median [Q1–Q3] and categorical variables as n (%). Neurological outcomes are presented hierarchically, with overall brain lesion status reported as the primary outcome and lesion subtypes described separately.

Nutritional management

All infants were managed according to a standardized institutional nutritional protocol applied consistently across the cohort. Parenteral nutrition was initiated immediately after birth with amino acids and 10% glucose. At approximately 24 hours of life, a multicomponent lipid emulsion (SMOFlipid; soybean oil, medium-chain triglycerides, olive oil, and fish oil) was introduced.

Minimal enteral feeding with maternal milk, or pasteurized donor milk when maternal milk was unavailable, was initiated within the first 24 hours of life. Upon reaching 80 mL/kg/day of enteral intake, human milk was fortified (Aptamil BMF) to provide additional protein and long-chain polyunsaturated fatty acids, including docosahexaenoic acid (DHA) and arachidonic acid (AA).

Parenteral lipid administration was discontinued when enteral intake reached 90-100 mL/kg/day. Amino acids and glucose infusion were adjusted according to standard clinical practice.

Plasma sample collection

Venous or capillary blood samples (0.5-1 mL) were collected longitudinally at five predefined timepoints and integrated with routine clinical care to minimize procedural burden. The first sample (T0) was obtained within 40 minutes after birth during placement of the umbilical venous catheter. Subsequent samples were collected at 48-72 hours of life (T1), 7 days of life (±1 day; T2), approximately 33 weeks postmenstrual age (T3), and at term-equivalent age (~40 weeks postmenstrual age; T4). When available, the umbilical venous catheter was used for early sampling; otherwise, peripheral venipuncture was performed. Samples at T3 and T4 were obtained via peripheral venipuncture.

Whole blood was collected into EDTA tubes and processed within 3 hours of collection. Plasma was isolated by centrifugation at 1200 × g for 15 minutes at room temperature. The supernatant was aliquoted, anonymized using alphanumeric identifiers, and stored at −80 °C until extracellular vesicle enrichment. Samples underwent a single freeze-thaw cycle for analysis.

For method optimization and technical validation experiments, anonymized pooled plasma from healthy adult donors was used. Individual plasma units were combined to generate homogeneous pooled material, which was aliquoted and processed in four independent technical replicates under identical experimental conditions. These pooled samples were used exclusively for workflow benchmarking and reproducibility assessment and were not included in longitudinal clinical analyses.

Automated extracellular vesicle enrichment and integrated proteomic-lipidomic sample preparation

EVs were enriched from plasma using a semi-automated SAX magnetic bead workflow adapted from the Mag-Net method[14] to enable parallel proteomic and lipidomic extraction from minimal

input volumes. All steps were performed on a KingFisher™ Apex magnetic handling station (Thermo Fisher Scientific) at 4 °C.

For each sample, 10 µL of plasma were thawed on ice and mixed with 40 µL PBS and 50 µL Bis-Tris Propane (BTP) binding buffer (100 mM BTP, pH 6.3, 150 mM NaCl), yielding a final volume of 100 µL. EV capture was performed using 10 µL of MagReSyn™ SAX magnetic beads (ReSyn Biosciences), pre-equilibrated in BTP wash buffer (50 mM BTP, pH 6.5, 150 mM NaCl). Sequential washes were performed to remove non-vesicular plasma components.

Bead-bound EVs were subjected to monophasic extraction using 600 µL 1-butanol/acetonitrile/water (3:1:1, v/v)[15]. This step promotes membrane disruption and the subsequent release of the molecular cargo. Samples were incubated overnight at −20 °C to induce protein precipitation and aggregation capture (PAC), while lipid species remained in the soluble organic phase.

The following day, samples were reprocessed on the KingFisher™ Apex platform. During the initial protein aggregation capture (PAC) step, the precipitated proteins were retained on the magnetic beads, while the lipid-rich supernatant was recovered and processed separately. The lipid fraction was clarified by centrifugation (4 °C, 4,000 × g, 25 min), transferred to clean tubes, vacuum-dried, and stored at −20 °C until LC-MS/MS analysis.

In parallel, the bead-bound proteins underwent automated processing, including sequential washes with acetonitrile (ACN), 70% ethanol, and isopropanol (IPA) to remove contaminants. On-bead digestion was performed in 100 mM Tris-HCl (pH 8.0) containing 10 mM tris(2-carboxyethyl)phosphine (TCEP) and 4 mM chloroacetamide (CAA) for reduction and alkylation. Proteins were digested at 37 °C for 4 h using LysC (Wako) and trypsin (Promega) at enzyme-to-protein ratios of 1:400 and 1:200 (w/w), respectively. Peptides were acidified with 10 µL of 2% (v/v) trifluoroacetic acid (TFA) and stored at −20 °C until LC-MS/MS analysis.

Data acquisition

Lipidomic LC-MS/MS acquisition. Lipidomic analyses were performed using a Vanquish Horizon UHPLC system coupled to an Orbitrap Fusion™ Tribrid™ mass spectrometer (Thermo Fisher Scientific). Dried lipid extracts were reconstituted in 1-butanol (BuOH)/isopropanol (IPA)/water (H_2O) (8:23:69, v/v) supplemented with 5 mM phosphoric acid, and 5 µL of each sample were injected onto a reversed-phase column.

Chromatographic separation was achieved on an ACQUITY UPLC BEH C18 column (2.1 × 50 mm, 1.7 µm; Waters) maintained at 65 °C, using a 7-minute gradient at a flow rate of 600 µL/min. Mobile phase A consisted of 10 mM ammonium formate in 60:40 ACN/H_2O with 0.1% formic acid (FA); mobile phase B consisted of 10 mM ammonium formate in 90:10 IPA/ACN with 0.1% FA. The gradient started at 15% B, increased to 30% B over 1 minute, then to 100% B over the following 5

minutes, and was held at 100% B for 1 minute. The column was subsequently re-equilibrated for 1 minute.

Eluting lipid species were analyzed by electrospray ionization tandem mass spectrometry (ESI-MS/MS). For technical validation experiments, data were acquired in both positive and negative ionization modes using a spray voltage of ±3.5 kV. For the longitudinal clinical dataset, acquisition was performed in positive ion mode only (+3.5 kV). The inlet capillary temperature was set to 300 °C. Nitrogen was used as sheath and auxiliary gas at flow rates of 30 and 10 arbitrary units (AU), respectively.

Data were acquired in data-dependent acquisition (DDA) mode, alternating full MS and MS/MS scans. Full MS survey scans were acquired in the Orbitrap analyzer over a mass range of 70-1000 m/z at 70,000 resolution (at 200 m/z), with an AGC target of 1 × 10□ and a maximum injection time of 100 ms. Up to five MS/MS events were triggered per MS scan using an isolation window of 1.4 Da and an intensity threshold of 1.6 × 10□. Fragment ion spectra were acquired at 17,500 resolution, with an AGC target of 1 × 10□ and a maximum injection time of 50 ms. Previously fragmented precursor ions were dynamically excluded for 2 seconds. Higher-energy collisional dissociation (HCD) was performed using stepped normalized collision energies of 20, 30, and 40.

Proteomic LC-MS/MS acquisition. Proteomic analysis was performed using an Evosep One chromatographic system coupled to an Orbitrap Exploris™ 480 mass spectrometer (Thermo Fisher Scientific) equipped with a FAIMS Pro Duo interface.

Peptide digests (20 µL) were loaded onto EvoTips according to the manufacturer's instructions. Peptide separation was achieved on an Aurora Rapid column (8 cm x 75 µm; IonOpticks) maintained at 25 °C, using the pre-programmed 60 samples per day (60 SPD) method at a constant flow rate of 1 µL/min.

Data were acquired in data-independent acquisition (DIA) mode. Full MS scans were acquired at 60,000 resolution over 380-980 m/z, with an automatic gain control (AGC) target of 300% and the maximum injection time (IT) set to Auto. MS/MS acquisition consisted of 50 DIA windows (15 Da width with 1 Da overlap) acquired at 15,000 resolution, with injection time set to Auto and a normalized collision energy of 25%. The FAIMS compensation voltage was set to −45 V at standard resolution. Data were acquired in profile mode using positive ion polarity.

Data processing and quality control

Lipidomic data processing. Raw lipidomics data were processed using MS-DIAL v5.5 for peak detection, deconvolution, retention time alignment, and lipid annotation[35]. The resulting master area table was refined using MS-FLO v1.8 to perform additional deisotoping and to resolve redundant or duplicated lipid features, thereby improving annotation consistency[36].

The curated table was imported into R v4.5.2 for downstream processing. Putative contaminants (e.g., chemical reagents, plasticizers, polymer-related compounds, and background signals) were excluded based on annotation keywords and pattern-based filtering.

Technical reproducibility was assessed using pooled quality control (QC) samples. For each lipid feature, mean intensity and coefficient of variation (CV) were calculated across QC injections. Features with CV >30% were removed.

To retain biologically informative species in the longitudinal dataset, a fill-rate filter was applied. Lipid features were retained if they exhibited ≥70% non-missing values in at least one timepoint. When multiple features corresponded to the same lipid annotation, a single representative feature was retained based on highest fill rate, lowest QC CV, and highest mean QC intensity.

Intensity values were $\log_2$-transformed and median-centered to correct for sample-wise technical variability. Missing values were imputed assuming left-censored missingness using a Gaussian distribution downshifted by 1.8 standard deviations with a width of 0.3 standard deviations.

Proteomic data processing. Proteomic data processing was performed separately for technical validation experiments and for the longitudinal clinical dataset using Spectronaut v18 (Biognosys AG)[37].

For technical validation experiments, raw data were analyzed using the Method Evaluation workflow to assess identification depth and analytical reproducibility.

For the longitudinal clinical dataset, raw files were processed in directDIA mode with default settings. Searches were performed against the UniProt human reference proteome with Trypsin/P and LysC specified as cleavage enzymes, carbamidomethylation (C) as a fixed modification, and methionine oxidation and protein N-terminal acetylation as variable modifications. False discovery rate was controlled at 1% at peptide spectrum match (PSM), peptide, and protein group levels using a target–decoy strategy. Protein inference and filtering were performed using the Identified (Q-value) setting, and protein quantification was based on MS2-level intensities.

Only the longitudinal dataset underwent downstream data curation in R (v4.5.2). Protein groups were harmonized to a single representative UniProt accession and gene symbol, prioritizing reviewed Swiss-Prot entries when available. Proteins were retained if they exhibited ≥70% non-missing values in at least one timepoint. Intensities ($\log_2$ scale) were median-centered across samples to correct for sample-wise technical variability.

Missing values in the longitudinal dataset were handled using a hybrid imputation strategy reflecting different plausible missingness mechanisms in MS-based proteomics.

As missing-at-random (MAR) and missing-not-at-random (MNAR) mechanisms cannot be directly identified from observed data alone, MNAR was operationally assumed when missingness patterns were consistent with left-censoring. Specifically, within each timepoint block, proteins exhibiting very low detection rates were classified as MNAR (proportion of observed values <0.30 when ≥12

samples were available in that timepoint, or <0.40 when <12 samples were available). These thresholds were selected to conservatively identify features with systematic under-detection while avoiding over-classification of sparse but potentially informative signals. Missing values not flagged as MNAR were treated as MAR. MAR values were imputed using structured least squares adaptive (SLSA) imputation (*imp4p* R package), whereas MNAR values were imputed from a downshifted Gaussian distribution (mean −1.8×SD, SD×0.3) to model signal intensities below the limit of detection.

Data analysis

All statistical analyses were performed in R v4.5.2. Proteomic and lipidomic datasets were analyzed separately and subsequently integrated using a network-based multi-omics framework.

Longitudinal molecular changes were assessed using per-feature linear mixed-effects models fitted with the *lmerTest* package. For each protein or lipid, abundance was modeled as a function of sampling timepoint (T0-T4) and gestational age at birth, with infant identity included as a random intercept to account for repeated measurements: abundance ~ timepoint + ga_birth + (1 | infant_id). Timepoint was treated as a categorical factor with T0 as the reference level. Models were fitted using maximum likelihood (REML = FALSE; optimizer = bobyqa).

A global time effect was assessed using Type-III ANOVA on the timepoint term, as implemented in *lmerTest*. For endpoint analyses, the contrast between term-equivalent age (T4) and birth (T0) was extracted from the model as the coefficient for timepoint T4. Effect sizes were reported as β coefficients, representing model-estimated differences in $\log_2$-transformed abundance. For timepoint effects, β corresponds to the $\log_2$ fold-change relative to the reference timepoint (T0), adjusted for gestational age at birth. P-values were adjusted using the Benjamini-Hochberg procedure. Features with FDR_global < 0.05 were considered significantly time-regulated. Features were considered significantly regulated at T4 versus T0 if they met FDR_T4vsT0 < 0.05 and |β| > 0.58, corresponding to an approximate 1.5-fold change on the original abundance scale, thereby prioritizing biologically meaningful shifts.

Cross-omics integration was performed using temporally regulated features (FDR_global < 0.05). Protein and lipid abundance profiles were z-score standardized, and pairwise Spearman correlations were computed across matched longitudinal samples. Correlation p-values were corrected using BH adjustment. Network edges were defined by FDR < 0.05 and |ρ| ≥ 0.35. The resulting weighted, undirected network was partitioned using Louvain community detection to identify cross-omics modules. Modules with fewer than 10 total features were excluded from downstream analyses.

Associations between molecular modules and clinical outcomes were tested separately for proteins and lipids using CAMERA gene-set testing within the *limma* framework. Models included the

clinical variable of interest as the main effect, with timepoint and gestational age at birth as covariates. Repeated measurements were accounted for using duplicateCorrelation with infant identity as the blocking factor. P-values were corrected for multiple testing across all module-outcome-layer combinations using the BH procedure. Associations with FDR < 0.05 were considered statistically significant.

Functional enrichment of proteomic clusters and network modules was performed in R using Gene Ontology Biological Process and Reactome pathway databases, using all quantified proteins as background and Benjamini–Hochberg correction (FDR < 0.05). Lipid ontology enrichment was performed using the LION web-based tool, ranking lipid species according to their T4 versus T0 β coefficients and retaining significantly enriched terms (FDR < 0.05).

Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository[38] with the dataset identifier PXD074778. *During peer review, the dataset can be accessed using the following credentials: Username: reviewer_pxd074778@ebi.ac.uk; Password: Ap7CxD6BMgE6; or via the access token: mMjU7VHYzPW2. These reviewer credentials will be removed prior to publication.*

Code availability

The custom R scripts used to generate the results reported in this study are available from the corresponding author upon request.

References

1. Ohuma, E. O. *et al.* National, regional, and global estimates of preterm birth in 2020, with trends from 2010: a systematic analysis. *The Lancet* **402**, 1261–1271 (2023).
2. Möllers, L. S. *et al.* Metabolic-endocrine disruption due to preterm birth impacts growth, body composition, and neonatal outcome. *Pediatr. Res.* **91**, 1350–1360 (2022).
3. Heath, R. J., Klevebro, S. & Wood, T. R. Maternal and Neonatal Polyunsaturated Fatty Acid Intake and Risk of Neurodevelopmental Impairment in Premature Infants. *Int. J. Mol. Sci.* **23**, 700 (2022).
4. Galinsky, R. *et al.* Complex interactions between hypoxia-ischemia and inflammation in preterm brain injury. *Dev. Med. Child Neurol.* **60**, 126–133 (2018).
5. Panfoli, I. *et al.* Oxidative Stress as a Primary Risk Factor for Brain Damage in Preterm Newborns. *Front. Pediatr.* **6**, 369 (2018).
6. Lee, A. H. *et al.* Dynamic molecular changes during the first week of human life follow a robust developmental trajectory. *Nat. Commun.* **10**, 1092 (2019).

7. Théry, C. *et al.* Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. *J. Extracell. Vesicles* **7**, 1535750 (2018).

8. Kalluri, R. & LeBleu, V. S. The biology, function, and biomedical applications of exosomes. *Science* **367**, eaau6977 (2020).

9. Young, K., Benny, M., Schmidt, A. & Wu, S. Unveiling the Emerging Role of Extracellular Vesicle–Inflammasomes in Hyperoxia-Induced Neonatal Lung and Brain Injury. *Cells* **13**, 2094 (2024).

10. Murphy, C. A. *et al.* Unique Patterns of Circulating Extracellular Vesicles in Preterm Infants During Adaptation to Extra-Uterine Life. *J. Extracell. Vesicles* **14**, e70064 (2025).

11. Skotland, T., Ekroos, K., Llorente, A. & Sandvig, K. Quantitative Lipid Analysis of Extracellular Vesicle Preparations: A Perspective. *J. Extracell. Vesicles* **14**, e70049 (2025).

12. Kong, F., Upadya, M., Weng, A. S. W., Dalan, R. & Dao, M. Isolating Small Extracellular Vesicles from Small Volumes of Blood Plasma using size exclusion chromatography and density gradient ultracentrifugation: A Comparative Study. Preprint at https://doi.org/10.1101/2023.10.30.564707 (2023).

13. Rai, A. *et al.* Multi-omics identify hallmark protein and lipid features of small extracellular vesicles circulating in human plasma. *Nat. Cell Biol.* 1–19 (2025) doi:10.1038/s41556-025-01795-7.

14. Wu, C. C. *et al.* Enrichment of extracellular vesicles using Mag-Net for the analysis of the plasma proteome. *Nat. Commun.* **16**, 5447 (2025).

15. Muehlbauer, L. K. *et al.* Rapid Multi-Omics Sample Preparation for Mass Spectrometry. *Anal. Chem.* **95**, 659–667 (2023).

16. Olin, A. *et al.* Stereotypic Immune System Development in Newborn Children. *Cell* **174**, 1277-1292.e14 (2018).

17. Zhong, W. *et al.* The development of blood protein profiles in extremely preterm infants follows a stereotypic evolution pattern. *Commun. Med.* **3**, 107 (2023).

18. Lozano-Andrés, E. *et al.* Physical association of low density lipoprotein particles and extracellular vesicles unveiled by single particle analysis. *J. Extracell. Vesicles* **12**, 12376 (2023).

19. Karimi, N. *et al.* Detailed analysis of the plasma extracellular vesicle proteome after separation from lipoproteins. *Cell. Mol. Life Sci.* **75**, 2873–2886 (2018).

20. Bennike, T. B. *et al.* Preparing for Life: Plasma Proteome Changes and Immune System Development During the First Week of Human Life. *Front. Immunol.* **11**, 578505 (2020).

21. Phuyal, S. *et al.* The Ether Lipid Precursor Hexadecylglycerol Stimulates the Release and Changes the Composition of Exosomes Derived from PC-3 Cells. *J. Biol. Chem.* **290**, 4225–4237 (2015).

22. Simbari, F. *et al.* Plasmalogen enrichment in exosomes secreted by a nematode parasite versus those derived from its mouse host: implications for exosome stability and biology. *J. Extracell. Vesicles* **5**, 30741 (2016).

23. Sódar, B. W. *et al.* Low-density lipoprotein mimics blood plasma-derived exosomes and microvesicles during isolation and detection. *Sci. Rep.* **6**, 24316 (2016).

24. Flaherty, S. E. *et al.* A lipase-independent pathway of lipid release and immune modulation by adipocytes. *Science* **363**, 989–993 (2019).

25. Huang, P. *et al.* Complement C3a induces axonal hypomyelination in the periventricular white matter through activation of WNT/β-catenin signal pathway in septic neonatal rats experimentally induced by lipopolysaccharide. *Brain Pathol.* **30**, 495–514 (2020).

26. Ten, V. S. *et al.* C1q-Deficiency Is Neuroprotective Against Hypoxic-Ischemic Brain Injury in Neonatal Mice. *Stroke* **36**, 2244–2250 (2005).

27. Ten, V. S. *et al.* Complement Component C1q Mediates Mitochondria-Driven Oxidative Stress in Neonatal Hypoxic–Ischemic Brain Injury. *J. Neurosci.* **30**, 2077–2087 (2010).

28. Pedroni, S. M. A. *et al.* Complement inhibition and statins prevent fetal brain cortical abnormalities in a mouse model of preterm birth. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.* **1842**, 107–115 (2014).

29. Liddelow, S. A. *et al.* Neurotoxic reactive astrocytes are induced by activated microglia. *Nature* **541**, 481–487 (2017).

30. Starke, N. *et al.* Extracellular Vesicle ASC: A Novel Mediator for Lung–Brain Axis in Preterm Brain Injury. *Am. J. Respir. Cell Mol. Biol.* **71**, 464–480 (2024).

31. Lal, C. V. *et al.* Exosomal microRNA predicts and protects against severe bronchopulmonary dysplasia in extremely premature infants. *JCI Insight* **3**, e93994 (2018).

32. Tóth, E. Á. *et al.* Formation of a protein corona on the surface of extracellular vesicles in blood plasma. *J. Extracell. Vesicles* **10**, e12140 (2021).

33. Dietz, L. *et al.* Uptake of extracellular vesicles into immune cells is enhanced by the protein corona. *J. Extracell. Vesicles* **12**, e12399 (2023).

34. Muraoka, S. *et al.* Comprehensive proteomic profiling of plasma and serum phosphatidylserine-positive extracellular vesicles reveals tissue-specific proteins. *iScience* **25**, 104012 (2022).

35. Tsugawa, H. *et al.* MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat. Methods* **12**, 523–526 (2015).

36. DeFelice, B. C. *et al.* Mass Spectral Feature List Optimizer (MS-FLO): A Tool To Minimize False Positive Peak Reports in Untargeted Liquid Chromatography–Mass Spectroscopy (LC-MS) Data Processing. *Anal. Chem.* **89**, 3250–3255 (2017).

37. Bruderer, R. *et al.* Extending the Limits of Quantitative Proteome Profiling with Data-Independent Acquisition and Application to Acetaminophen-Treated Three-Dimensional Liver Microtissues. *Mol. Cell. Proteomics* **14**, 1400–1410 (2015).

38. Perez-Riverol, Y. *et al.* The PRIDE database at 20 years: 2025 update. *Nucleic Acids Res.* **53**, D543–D553 (2025).

Author contributions

C.L. and A.P. conceived and designed the study. N.G. and F.A. performed the experiments and curated the dataset. N.G. optimized the workflow, conducted the bioinformatic and statistical analyses, and drafted the manuscript. F.A. and M.Ba. contributed to data interpretation. L.R. and S.S. contributed to sample processing and data acquisition. G.T., C.A., F.V., and L.A.R. contributed to patient recruitment, clinical data collection, and clinical interpretation. G.L. and M.Bru. contributed to data interpretation and critically revised the manuscript. L.A.R., C.L., and A.P. jointly supervised the study. All authors reviewed and approved the manuscript.

Competing interests

The authors declare no competing interests.

Figure legends

Fig.1: Technical validation of the EV-based multi-omics workflow using plasma micro-volumes. (a) Mean peptide and protein group identifications obtained from 5 µL and 10 µL plasma inputs following EV enrichment and monophasic solvent extraction with 1-butanol/acetonitrile/water (3:1:1, v/v; BAMM), isopropanol (IPA), or methanol (MeOH), compared to the standard Mag-Net protocol. Dots represent technical replicates; bars indicate mean values. “Neat” denotes direct plasma digestion without EV enrichment. (b) Distribution of coefficients of variation (CV) for quantified protein groups across conditions. Bars indicate the proportion of proteins within predefined CV ranges. (c) Scatter plots of $\log_2$-transformed protein intensities comparing each extraction condition with the corresponding standard Mag-Net reference at matched plasma input. Pearson correlation coefficients (r) are shown; dashed lines indicate the identity line. (d) Log_2 fold change (EV/Neat) is shown for representative enriched EV core markers and depleted abundant plasma proteins under BAMM 10 µL and Mag-Net standard 10 µL conditions. (e) Lipidomic depth after application of an intensity threshold (>1×10□). Mean numbers of detected lipid species in negative and positive ionization modes across extraction strategies and plasma input volumes. Dots represent technical replicates; bars indicate mean values. QC denotes pooled quality control samples. (f) Distribution of CV (%) for quantified lipid species in negative and positive polarity. Dashed vertical lines indicate 20% and 30% CV thresholds.

Fig. 2: Longitudinal study design and plasma EV multi-omics profiling in very preterm infants. (a) Overview of the longitudinal cohort, sampling timeline, and analytical workflow. Plasma samples were collected from very preterm infants born at <32 weeks of gestational age (n = 16) at five timepoints spanning birth to term-equivalent age: birth (T0), 48-72 hours (T1), 7 days (T2), 33 weeks postmenstrual age (T3), and 40 weeks postmenstrual age (T4, n = 10). EVs were enriched from plasma, followed by parallel isolation of EV-associated proteins and lipids from the same sample. Proteomic and lipidomic profiles were generated by LC-MS/MS and processed through standardized pipelines. (b) Dynamic range of the plasma EV proteome shown as a rank-abundance plot of median $\log_2$ protein abundance across all samples, highlighting representative EV-associated proteins. (c) Subcellular localization enrichment analysis of the EV proteome based on the Jensen COMPARTMENTS database. Terms are ordered by associated protein count along the x axis. Dot size reflects statistical significance ($-\log_{10}$ FDR), while color intensity indicates odds ratio. (d) Distribution of identified lipid species across major lipid categories, including glycerophospholipids, glycerolipids, sphingolipids, fatty acyls, and sterol lipids.

Fig. 3: Coordinated longitudinal remodeling of the plasma EV proteome from birth to term-equivalent age. (a) Circular heatmap illustrating the 875 EV proteins with a significant temporal profile (LMM; FDR < 0.05). Data are shown as row z-scores of median protein abundances across five postnatal timepoints, from birth (T0) to term-equivalent age (TEA, T4). Hierarchical clustering resolved four primary modules; the outer ring indicates the direction of the net developmental shift (T4 vs T0). (b) Cluster-specific temporal trajectories. Profile plots show z-scored abundances for individual proteins (thin lines) and cluster medians (bold lines), highlighting coordinated decreasing (clusters 1A, 1B) and increasing (clusters 2A, 2B) proteomic programs (n indicates protein count per cluster). (c) Volcano plot displaying the differential protein abundance between TEA (T4) and birth (T0) derived from the LMM. Statistical significance (−log10 FDR) is plotted against the effect size (β coefficient). Significant features (FDR < 0.05 and |β| > 0.58) are highlighted, with counts of downregulated (blue, 439) and upregulated (red, 127) proteins indicated. (d) Longitudinal profiles of representative EV-associated proteins. Boxes represent median and interquartile range; whiskers indicate data range. (e) Functional enrichment analysis of significantly downregulated (left) or upregulated proteins (right) at TEA. Top enriched terms from Gene Ontology (Biological Process) and Reactome databases are ranked by statistical significance (−log10 FDR). Dot size reflects the number of proteins associated with each term.

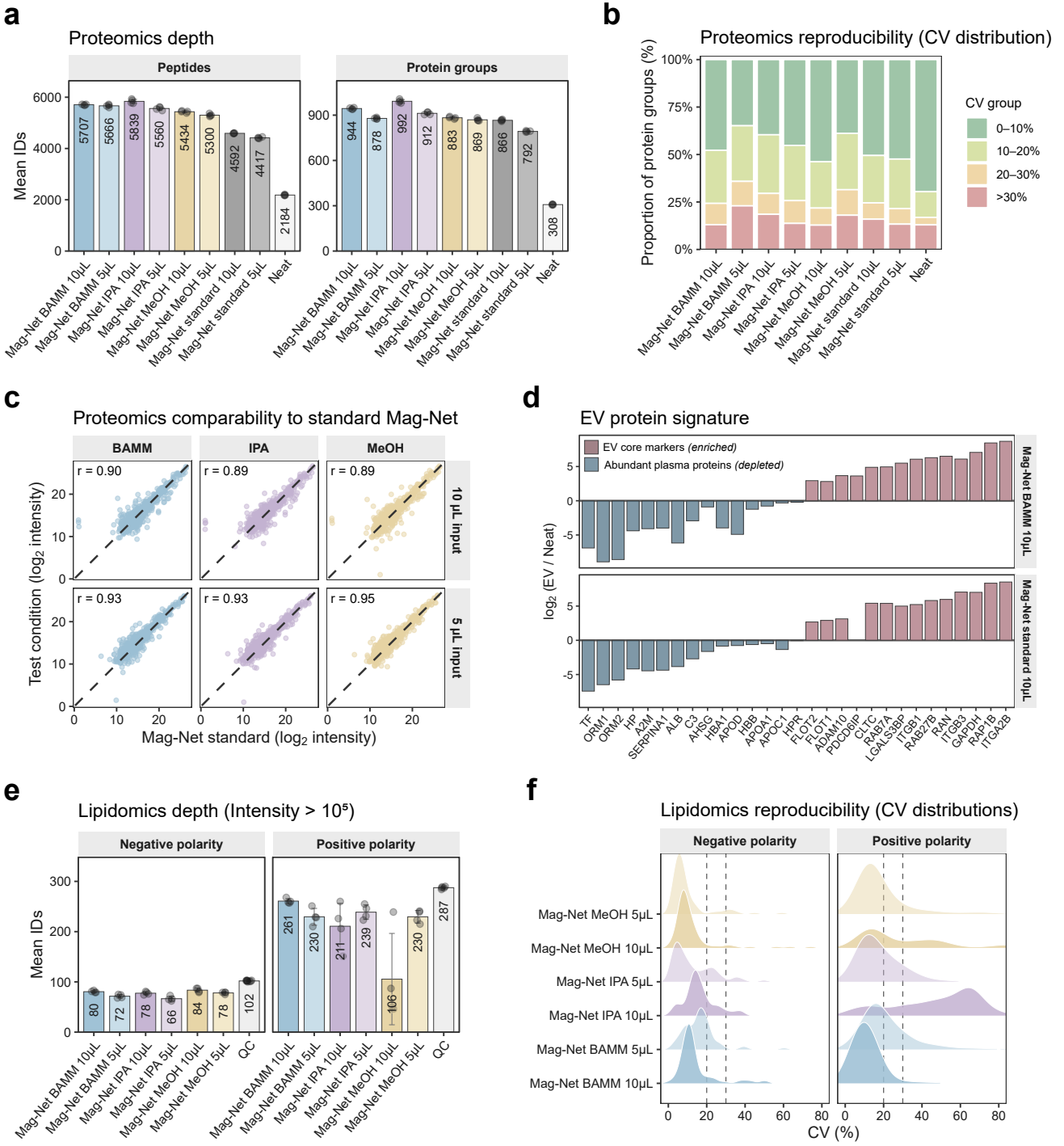
Fig. 4: Coordinated structural and metabolic maturation of the plasma EV lipidome. (a) Heatmap of the 100 lipid species showing a significant temporal profile in the longitudinal linear mixed-effects model (LMM; FDR < 0.05). Data are represented as row z-scores of median abundances across five postnatal timepoints (T0-T4). Hierarchical clustering identifies four distinct

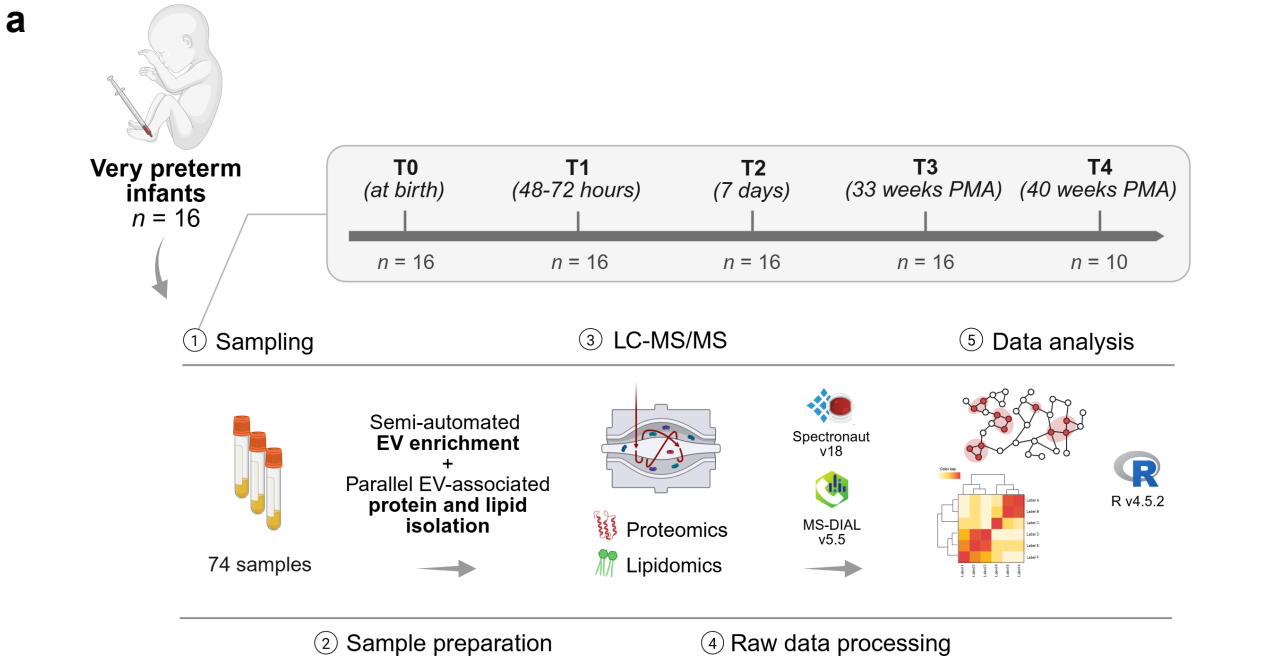
longitudinal trajectories; the right-side annotation indicates the direction of change at term-equivalent age (TEA, T4) relative to birth (T0). (b) Cluster-specific temporal trajectories and compositional analysis. Profile plots (left) show z-scored abundances for individual lipids (thin lines) and cluster medians (bold lines). The distribution of lipid classes within each cluster is shown as pie charts (right). (c) Volcano plot displaying differential lipid abundance between TEA (T4) and birth (T0). Statistical significance (−log10 FDR) is plotted against the effect size (β coefficient). Significant features (FDR < 0.05, |β| > 0.58) are highlighted, with counts of upregulated (red, 41) and downregulated (green, 22) lipids indicated. (d) Lipid ontology (LION) enrichment analysis based on the β coefficient (T4 vs T0). Enrichment scores are represented as lollipop plots, where dot size reflects the number of annotated lipid species and color intensity indicates statistical significance (−log10 FDR). (e) Structural space distribution of regulated phosphatidylcholines (PC), ether-linked phosphatidylcholines (PC O-), and sphingomyelins (SM). Lipids are mapped according to the number of carbons in the side chain (x-axis) and degree of unsaturation (y-axis). Dot size reflects statistical significance (−log10 FDR) and color encodes effect size (T4 vs T0), highlighting selective structural redistribution across postnatal development.

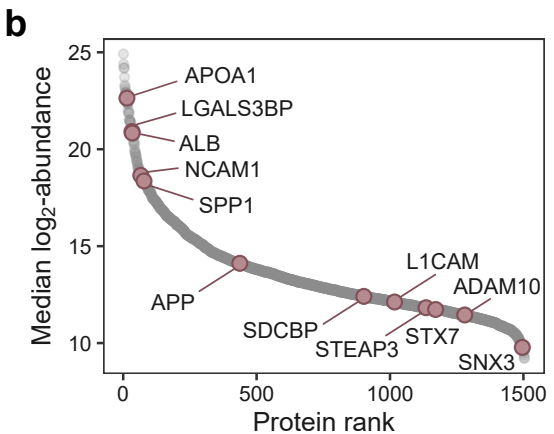
Fig. 5: Cross-omics integration of the plasma EV proteome and lipidome. (a) Schematic overview of the cross-omics integration strategy. Temporally regulated EV proteins (n = 875) and lipids (n = 100), identified by longitudinal linear mixed-effects modeling (LMM), were integrated using a protein-lipid correlation network based on Spearman correlations across paired longitudinal samples. Louvain community detection was applied to identify cross-omics modules, which were subsequently tested for associations with clinical outcomes. (b) Distribution of Spearman correlation coefficients (ρ) for all protein-lipid pairs. Only significant edges (FDR < 0.05 and |ρ| ≥ 0.35) were retained for network construction. (c) Composition of the five cross-omics modules (M1-M5) by molecular layer (protein vs lipid). Bars indicate the relative contribution of proteins and lipids within each module, with total node counts reported above each bar. (d) Functional enrichment analysis of the protein components within each module based on the Gene Ontology Biological Process database. Terms are ranked by −log10(FDR) and colored by module assignment.

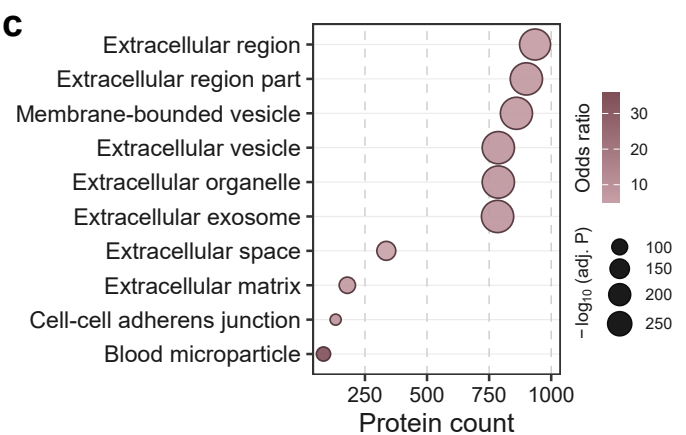
Fig. 6: Integrated EV modules show significant associations with clinical phenotypes in preterm infants. (a) Heatmap summarizing module-level associations between EV molecular modules and clinical outcomes, assessed using a CAMERA-based competitive gene set framework. Protein (top) and lipid (bottom) groups within each module were tested separately. Colors represent signed −log10(p-values), with red indicating positive associations and blue negative associations; asterisks denote associations significant after global FDR correction. Any brain lesion includes both hemorrhagic brain injury and white matter injury. While only a limited

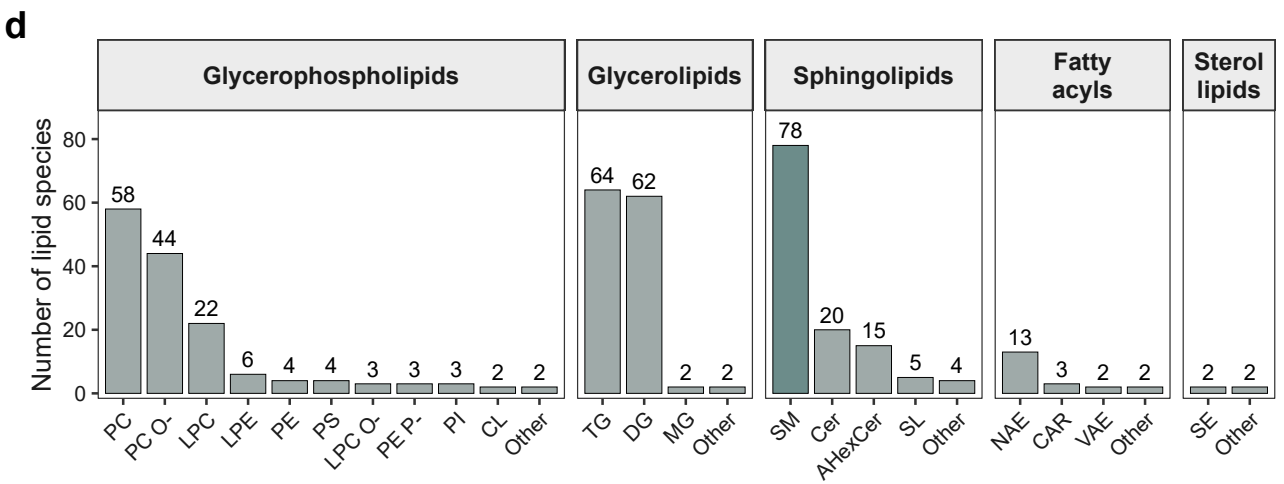
number of individual lipid species reached statistical significance after multiple testing, coordinated shifts across lipid modules were detected at the module level, highlighting the ability of module-based approaches to capture distributed but consistent EV molecular patterns. (b–d) Longitudinal trajectories of representative leading-edge candidates illustrating the molecular programs underlying significant module–outcome associations. Lines represent mean ± s.e.m. across timepoints (T0–T4).

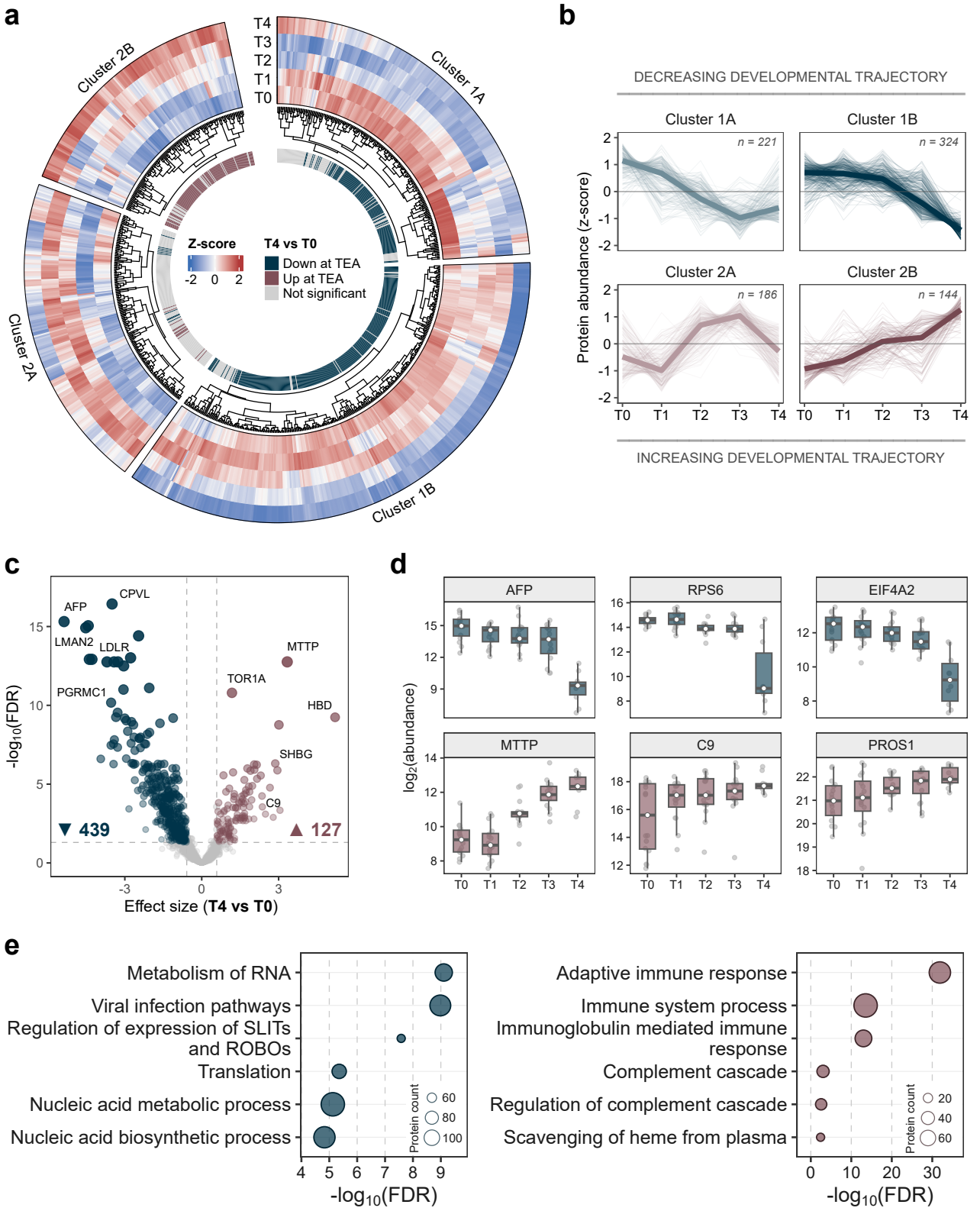

a
Proteomics depth
Peptides
Protein groups
Mean IDs
5707
5666
5839
5560
5434
5300
4592
4417
2184
944
878
992
912
883
869
866
792
308
Mag-Net BAMM 10µL
Mag-Net BAMM 5µL
Mag-Net IPA 10µL
Mag-Net IPA 5µL
Mag-Net MeOH 10µL
Mag-Net MeOH 5µL
Mag-Net standard 10µL
Mag-Net standard 5µL
Neat
b
Proteomics reproducibility (CV distribution)
Proportion of protein groups (%)
CV group
0–10%
10–20%
20–30%
>30%
c
Proteomics comparability to standard Mag-Net
BAMM
IPA
MeOH
10 µL input
5 µL input
r = 0.90
r = 0.89
r = 0.89
r = 0.93
r = 0.93
r = 0.95
Test condition (log2 intensity)
Mag-Net standard (log2 intensity)
d
EV protein signature
EV core markers (enriched)
Abundant plasma proteins (depleted)
log2 (EV / Neat)
Mag-Net BAMM 10µL
Mag-Net standard 10µL
TF
ORM1
ORM2
HP
A2M
SERPINA1
ALB
C3
AHSG
HBA1
APOD
HBB
APOA1
APOC1
HPR
FLOT2
FLOT1
ADAM10
PDCD6IP
CLTC
RAB7A
LGALS3BP
ITGB1
RAB27B
RAN
ITGB3
GAPDH
RAP1B
ITGA2B
e
Lipidomics depth (Intensity > 10⁵)
Negative polarity
Positive polarity
Mean IDs
80
72
78
66
84
78
102
261
230
211
239
106
230
287
QC
f
Lipidomics reproducibility (CV distributions)
Mag-Net MeOH 5µL
Mag-Net MeOH 10µL
Mag-Net IPA 5µL
Mag-Net IPA 10µL
Mag-Net BAMM 5µL
Mag-Net BAMM 10µL
CV (%)

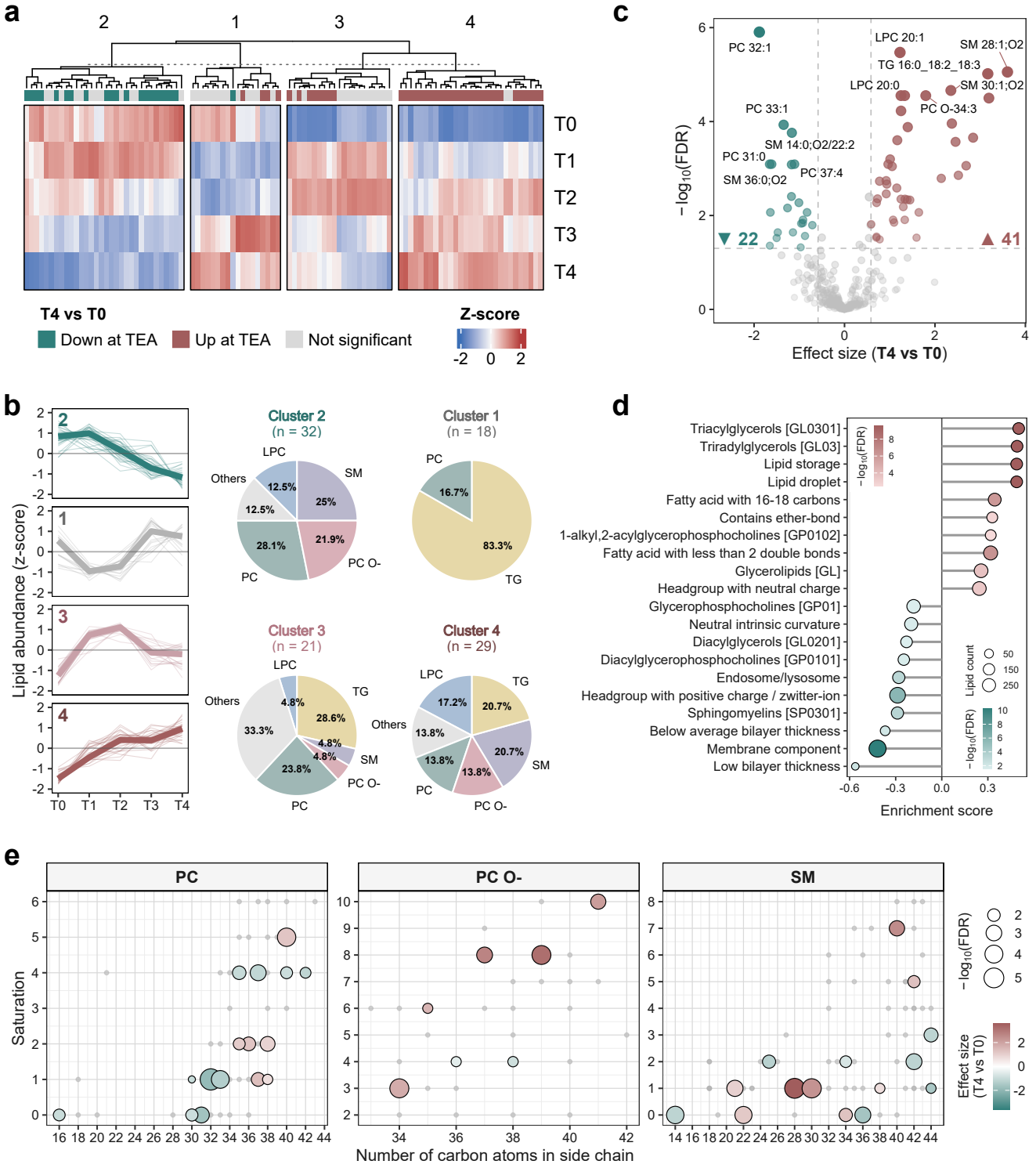

a
Very preterm infants
n = 16
T0 (at birth)
T1 (48-72 hours)
T2 (7 days)
T3 (33 weeks PMA)
T4 (40 weeks PMA)
n = 16
n = 16
n = 16
n = 16
n = 10
① Sampling
③ LC-MS/MS
⑤ Data analysis
74 samples
Semi-automated EV enrichment + Parallel EV-associated protein and lipid isolation
Proteomics
Lipidomics
Spectronaut v18
MS-DIAL v5.5
R v4.5.2
② Sample preparation
④ Raw data processing

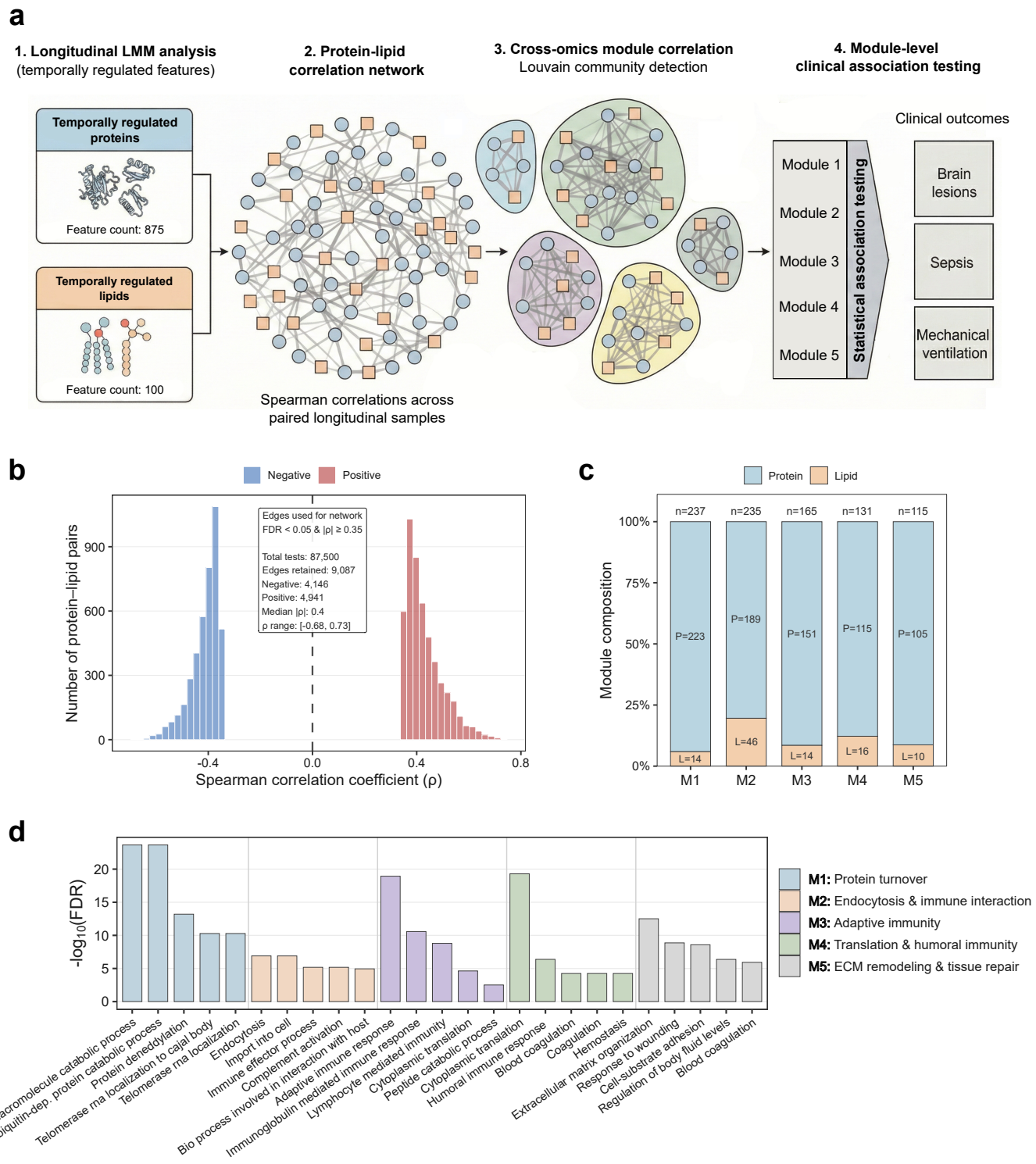

b
Median log2-abundance
Protein rank
APOA1
LGALS3BP
ALB
NCAM1
SPP1
APP
L1CAM
ADAM10
SDCBP
STEAP3
STX7
SNX3

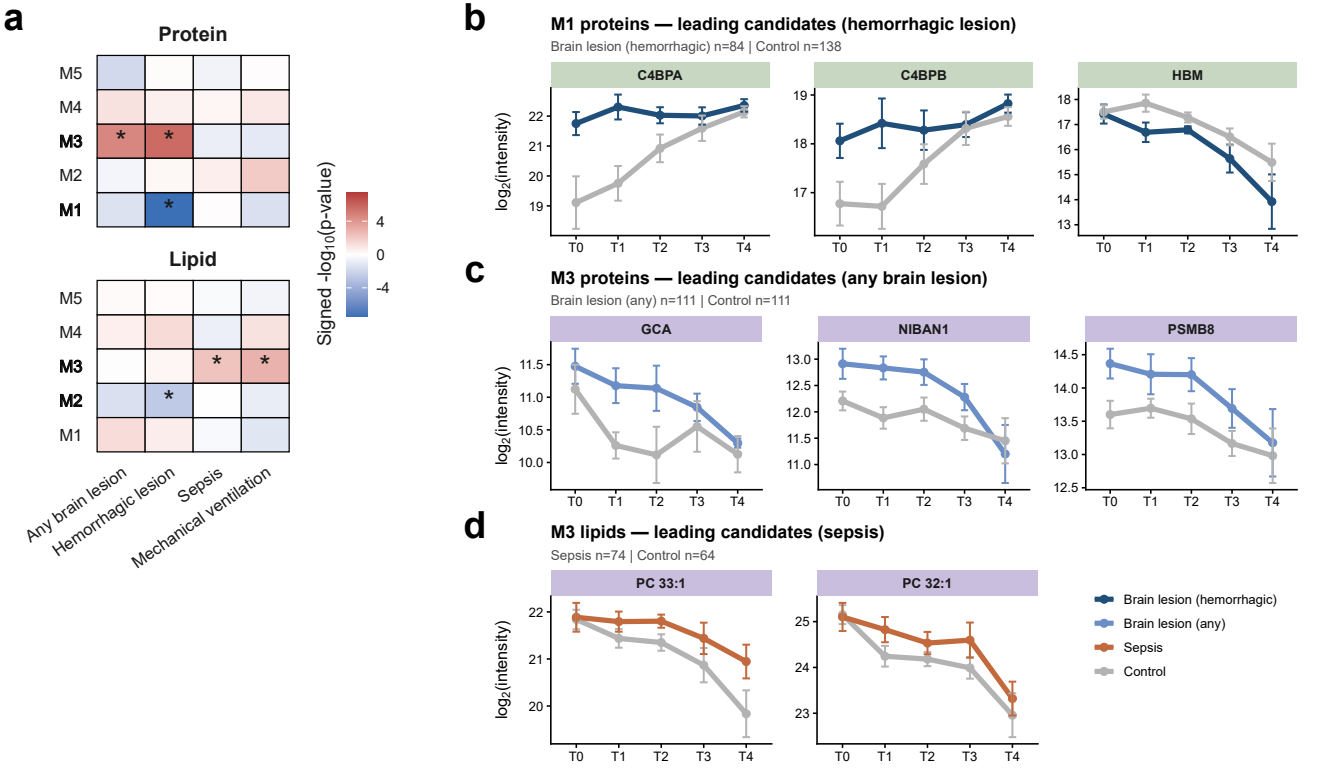

c
Extracellular region
Extracellular region part
Membrane-bounded vesicle
Extracellular vesicle
Extracellular organelle
Extracellular exosome
Extracellular space
Extracellular matrix
Cell-cell adherens junction
Blood microparticle
Protein count
Odds ratio
−log10 (adj. P)


d
Number of lipid species
Glycerophospholipids
Glycerolipids
Sphingolipids
Fatty acyls
Sterol lipids
PC
PC O-
LPC
LPE
PE
PS
LPC O-
PE P-
PI
CL
Other
TG
DG
MG
Other
SM
Cer
AHexCer
SL
Other
NAE
CAR
VAE
Other
SE
Other


a
T4
T3
T2
T1
T0
Cluster 1A
Cluster 1B
Cluster 2A
Cluster 2B
Z-score
-2 0 2
T4 vs T0
Down at TEA
Up at TEA
Not significant
b
DECREASING DEVELOPMENTAL TRAJECTORY
Cluster 1A
Cluster 1B
Cluster 2A
Cluster 2B
n = 221
n = 324
n = 186
n = 144
Protein abundance (z-score)
T0 T1 T2 T3 T4
INCREASING DEVELOPMENTAL TRAJECTORY
c
AFP
CPVL
LMAN2
LDLR
PGRMC1
MTTP
TOR1A
HBD
SHBG
C9
▼ 439
▲ 127
-log10(FDR)
Effect size (T4 vs T0)
d
AFP
RPS6
EIF4A2
MTTP
C9
PROS1
log2(abundance)
T0 T1 T2 T3 T4
e
Metabolism of RNA
Viral infection pathways
Regulation of expression of SLITs and ROBOs
Translation
Nucleic acid metabolic process
Nucleic acid biosynthetic process
Protein count
60
80
100
-log10(FDR)
Adaptive immune response
Immune system process
Immunoglobulin mediated immune response
Complement cascade
Regulation of complement cascade
Scavenging of heme from plasma
Protein count
20
40
60
-log10(FDR)


a
2
1
3
4
T0
T1
T2
T3
T4
T4 vs T0
Down at TEA
Up at TEA
Not significant
Z-score
-2 0 2
c
PC 32:1
LPC 20:1
SM 28:1;O2
TG 16:0_18:2_18:3
LPC 20:0
SM 30:1;O2
PC O-34:3
PC 33:1
SM 14:0;O2/22:2
PC 31:0
PC 37:4
SM 36:0;O2
▼ 22
▲ 41
−log10(FDR)
Effect size (T4 vs T0)
b
Lipid abundance (z-score)
Cluster 2
(n = 32)
LPC
SM
Others
12.5%
25%
12.5%
28.1%
21.9%
PC O-
PC
Cluster 1
(n = 18)
PC
16.7%
83.3%
TG
Cluster 3
(n = 21)
LPC
TG
Others
4.8%
28.6%
33.3%
4.8%
4.8%
SM
23.8%
PC O-
PC
Cluster 4
(n = 29)
LPC
TG
17.2%
20.7%
Others
13.8%
20.7%
SM
13.8%
13.8%
PC
PC O-
T0 T1 T2 T3 T4
d
Triacylglycerols [GL0301]
Triradylglycerols [GL03]
Lipid storage
Lipid droplet
Fatty acid with 16-18 carbons
Contains ether-bond
1-alkyl,2-acylglycerophosphocholines [GP0102]
Fatty acid with less than 2 double bonds
Glycerolipids [GL]
Headgroup with neutral charge
Glycerophosphocholines [GP01]
Neutral intrinsic curvature
Diacylglycerols [GL0201]
Diacylglycerophosphocholines [GP0101]
Endosome/lysosome
Headgroup with positive charge / zwitter-ion
Sphingomyelins [SP0301]
Below average bilayer thickness
Membrane component
Low bilayer thickness
−log10(FDR)
Lipid count
Enrichment score
e
PC
PC O-
SM
Saturation
Number of carbon atoms in side chain
−log10(FDR)
Effect size (T4 vs T0)


a
1. Longitudinal LMM analysis
(temporally regulated features)
2. Protein-lipid correlation network
3. Cross-omics module correlation
Louvain community detection
4. Module-level clinical association testing
Temporally regulated proteins
Feature count: 875
Temporally regulated lipids
Feature count: 100
Spearman correlations across paired longitudinal samples
Module 1
Module 2
Module 3
Module 4
Module 5
Statistical association testing
Clinical outcomes
Brain lesions
Sepsis
Mechanical ventilation
b
Negative
Positive
Edges used for network
FDR < 0.05 & |ρ| ≥ 0.35
Total tests: 87,500
Edges retained: 9,087
Negative: 4,146
Positive: 4,941
Median |ρ|: 0.4
ρ range: [-0.68, 0.73]
Number of protein–lipid pairs
Spearman correlation coefficient (ρ)
c
Protein
Lipid
Module composition
n=237
n=235
n=165
n=131
n=115
P=223
P=189
P=151
P=115
P=105
L=14
L=46
L=14
L=16
L=10
M1
M2
M3
M4
M5
d
-log10(FDR)
M1: Protein turnover
M2: Endocytosis & immune interaction
M3: Adaptive immunity
M4: Translation & humoral immunity
M5: ECM remodeling & tissue repair
Macromolecule catabolic process
Ubiquitin-dep. protein catabolic process
Protein deneddylation
Telomerase rna localization to cajal body
Telomerase rna localization
Endocytosis
Import into cell
Immune effector process
Complement activation
Bio process involved in interaction with host
Adaptive immune response
Immunoglobulin mediated immune response
Lymphocyte mediated immunity
Cytoplasmic translation
Peptide catabolic process
Cytoplasmic translation
Humoral immune response
Blood coagulation
Coagulation
Hemostasis
Extracellular matrix organization
Response to wounding
Cell-substrate adhesion
Regulation of body fluid levels
Blood coagulation


a
Protein
M5
M4
M3
M2
M1
Lipid
M5
M4
M3
M2
M1
Any brain lesion
Hemorrhagic lesion
Sepsis
Mechanical ventilation
Signed -log10(p-value)
4
0
-4
b
M1 proteins — leading candidates (hemorrhagic lesion)
Brain lesion (hemorrhagic) n=84 | Control n=138
C4BPA
C4BPB
HBM
log2(intensity)
T0
T1
T2
T3
T4
c
M3 proteins — leading candidates (any brain lesion)
Brain lesion (any) n=111 | Control n=111
GCA
NIBAN1
PSMB8
log2(intensity)
d
M3 lipids — leading candidates (sepsis)
Sepsis n=74 | Control n=64
PC 33:1
PC 32:1
log2(intensity)
Brain lesion (hemorrhagic)
Brain lesion (any)
Sepsis
Control