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

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

[1]Core Facility for Omics Sciences, IRCCS Istituto Giannina Gaslini, Genoa, Italy.
[2]Department of Experimental Medicine (DIMES), University of Genoa, Italy.
[3]Division of Nephrology, Dialysis and Transplantation and Laboratory of Molecular Nephrology, Istituto Giannina Gaslini, IRCCS, Genoa, Italy.
[4]Transfusion Medicine Department, IRCCS Ospedale Policlinico San Martino, Genoa, Italy.
[5]Neonatal Intensive Care Unit, Department of Maternal and Neonatal Health, IRCCS Istituto Giannina Gaslini, 16147, Genoa, Italy.
[6]Neonatal Intensive Care Unit, Department of Maternal and Neonatal Health, IRCCS Istituto Giannina Gaslini, 16147, Genoa, Italy; Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Mother and Child Health, School of Medical and Pharmaceuticals, University of Genoa, 16132, Genoa, Italy.

*These authors contributed equally to this work. # These authors jointly supervised this work. & orresponding authors: chiaralavarello@gaslini.org

Abstract

Introduction. Very preterm birth (<32 weeks of gestation) disrupts key developmental programs and exposes infants to a highly vulnerable postnatal period. However, the molecular trajectories underlying early postnatal adaptation remain poorly characterized, largely due to limitations in longitudinal sampling and the need for minimally invasive approaches.

Methods. Here, we applied an integrated longitudinal proteomic and lipidomic analysis to plasma extracellular vesicles (EVs) from very preterm infants, spanning the developmental window from birth to term-equivalent age. EVs were enriched

from plasma micro-volumes (10 µL) using an automated Mag-Net-based workflow and analyzed by high-resolution mass spectrometry. Longitudinal molecular changes were modeled using linear mixed-effects models, and cross-omics integration was performed through correlation network analysis and module detection. Associations with clinical outcomes were assessed using a module-level competitive testing framework.

Results. We quantified 1,528 EV-associated proteins and 421 lipid species across 74 longitudinal samples. The EV proteome underwent extensive and structured remodeling over time, characterized by a progressive decrease in proteins related to translation and early growth processes and a concomitant increase in immune-related components, including complement and immunoglobulin-associated proteins. In contrast, lipidomic remodeling involved a more selective subset of species but followed coordinated temporal patterns. Lipids increasing toward term-equivalent age were enriched for triacylglycerols and ether-linked phosphatidylcholines, while specific membrane-associated species decreased, indicating selective structural reorganization rather than global class-level shifts.
Cross-omics integration revealed five coordinated protein-lipid modules, each capturing coherent longitudinal programs. Several modules showed significant associations with clinical features of prematurity, particularly brain injury, highlighting that EV molecular trajectories reflect clinically relevant biological states. Module-based associations were more robust than single-feature signals, especially for lipidomic data.

Conclusions. Together, these results demonstrate that EV-based multi-omics enables longitudinal molecular profiling in very preterm infants using minimal sample volumes and captures coordinated biological programs linked to development and clinical outcomes. This framework provides a scalable and biologically informative approach for studying neonatal adaptation and vulnerability in fragile clinical populations.

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

Conflict of interest statement

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Andrea Petretto reports financial support was provided by Giannina Gaslini Institute. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

1. Introduction

1.1 Preterm birth and early postnatal vulnerability

Preterm birth affects approximately 13-15 million infants each year worldwide (≈1 in 10 births) and is consistently identified as the leading cause of neonatal and under-five mortality and a major source of long-term neurodevelopmental disability[1]. Despite substantial advances in neonatal intensive care and improved survival across gestational ages, infants born very preterm (<32 weeks of gestation)

face a biological condition that is qualitatively distinct from both term birth and later prematurity.

Very preterm birth is not simply an earlier transition to extrauterine life, but rather an abrupt disruption of the maternal-placental-fetal unit during a critical window of physiological maturation[2]. The third trimester of gestation is characterized by rapid somatic growth and tightly regulated developmental programs, including the placental transfer of long-chain polyunsaturated fatty acids, such as docosahexaenoic acid and arachidonic acid, which are essential for brain development, adipose tissue accretion, and metabolic homeostasis[2,3]. In parallel, this period involves the transfer of maternal hormones, growth factors, and immunoglobulins that support metabolic, endocrine, and immune maturation.
Interruption of these processes by very preterm birth results in a state of metabolic imbalance and profound immune immaturity[4]. In this context, limited cerebral autoregulation and immature antioxidant defenses make the early postnatal period particularly vulnerable to external stressors, including hyperoxia, infection, and inflammation[5]. These early exposures drive oxidative and inflammatory cascades that interfere with myelination, neuronal maturation, and vascular stability, with consequences that extend well beyond the neonatal period and contribute to long-term neurodevelopmental risk[6,7].

1.2 Limitations of current approaches and the need for longitudinal molecular profiling

Clinical monitoring in neonatal intensive care units is largely based on physiological parameters and population-level reference ranges. While essential for acute management, these measures are poorly suited to capture the rapid and heterogeneous molecular changes that characterize early postnatal adaptation in very preterm infants. At the molecular level, most studies rely on cross-sectional designs that compare infants at different postnatal ages or clinical states. Such approaches are particularly vulnerable to inter-individual variability, which is pronounced in very preterm cohorts due to differences in gestational age, clinical exposures, and disease burden.

Longitudinal study designs, in which each infant serves as their own biological reference over time, are therefore essential to disentangle intrinsic developmental programs from reactive molecular responses associated with clinical stress or injury. This distinction is especially critical in the neonatal setting, where developmental timing and pathological exposures are tightly intertwined during the first weeks of life[8]. Systematic longitudinal molecular profiling provides a framework to capture coordinated trajectories of postnatal maturation and to identify early deviations that may precede or accompany adverse clinical outcomes.

However, the implementation of longitudinal omics studies in very preterm infants is constrained by limited sample volumes, ethical considerations, and the need for minimally invasive sampling strategies. These constraints have historically limited the depth and integration of molecular profiling in this population.

1.3 Extracellular vesicles as a platform for longitudinal multi-omics in very preterm infants

In this framework, extracellular vesicles (EVs) represent a promising system-level readout of neonatal biology. EVs are membrane-bound nanoparticles released by virtually all cell types and carry a protected cargo of proteins and lipids that reflects the physiological state of their cells of origin[9,10]. Through their role in intercellular communication, EVs are involved in processes relevant to early development, including immune regulation, metabolic adaptation, and neurodevelopmental signaling.

From a practical perspective, EVs can be isolated from plasma using minimally invasive blood sampling, making them particularly well suited for longitudinal studies in very preterm infants[11]. Importantly, EVs provide access to multiple molecular layers from the same biological compartment, enabling integrated proteomic and lipidomic analyses that capture both functional and structural aspects of postnatal adaptation[12].

Despite this potential, the application of EV-based liquid biopsy in very preterm infants has been limited by technical constraints related to EV isolation. Most

established EV enrichment strategies, including ultracentrifugation, density-gradient separation, and size-based methods, require relatively large plasma volumes and multiple processing steps. These requirements are poorly compatible with neonatal studies and make dense longitudinal sampling difficult. Moreover, incomplete separation of EVs from abundant non-vesicular plasma components, particularly lipoproteins, represents a major obstacle for EV lipidomics and can obscure EV-specific molecular signals[13].

To address these limitations, the present study builds on an automated strong anion-exchange-based enrichment workflow (Mag-Net[14]), originally developed and validated for plasma proteomics through the selective enrichment of extracellular vesicles. By exploiting the net negative surface charge of EV membranes, this approach enables efficient isolation of vesicle-associated material from complex plasma matrices while minimizing contamination from abundant soluble proteins.
Here, the Mag-Net workflow is adapted to enable integrated EV proteomic and lipidomic profiling from plasma micro-volumes (10 µL). By coupling EV enrichment with a monophasic extraction strategy, proteins and lipids are recovered from the same EV preparation, preserving the native coupling between vesicle structure and function. This framework establishes a longitudinal EV multi-omics approach whose feasibility and analytical potential are demonstrated in a clinically characterized cohort of very preterm infants – a uniquely informative setting in which rapid postnatal maturation is expected to be reflected in coordinated remodeling of EV protein and lipid composition.

3. Materials and methods

3.1 Cohort description

This study was conducted on a prospective, longitudinal cohort comprising 16 very preterm infants admitted to the Neonatal Pathology and Intensive Care Unit of IRCCS Istituto Giannina Gaslini (Genoa, Italy). Infants were consecutively enrolled between November 2022 and July 2024 and followed from birth to term-equivalent age according to the study protocol.

The cohort size reflects the clinical and ethical constraints typical of neonatal research, particularly in very preterm infants. For this reason, a longitudinal study design was adopted to maximize the information obtained from each subject. Each infant contributed samples at multiple postnatal timepoints, resulting in a total of 74 plasma samples per omics layer, enabling the investigation of molecular changes over time within the same individuals rather than relying solely on cross-sectional comparisons.

Demographic and clinical data were collected from medical records. A limited set of baseline variables was selected based on relevance to neonatal maturation and data availability, in order to reduce confounding and maintain interpretability in a small cohort. An overview of the baseline demographic and clinical characteristics of the study population is provided in **Table 1**.

Neurological status was assessed using routine clinical neuroimaging, including cranial ultrasound and magnetic resonance imaging. Brain injury was recorded as a binary variable indicating the presence or absence of any brain lesion. When present, lesions were further classified as hemorrhagic or white matter lesions for descriptive purposes. In order to account for major clinical conditions that could influence the plasma molecular profile, selected variables were recorded for each infant. In particular, sepsis – defined as a positive blood culture or antibiotic therapy lasting at least five days – and mechanical ventilation were included as indicators of systemic inflammatory and respiratory stress during the neonatal period.

The study was conducted in accordance with the Declaration of Helsinki and was 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: Characteristics of the study cohort.

Baseline demographic and clinical characteristics of very preterm infants included in the study (N = 16). Continuous variables are reported as *median [Q1–Q3]*, while categorical variables are reported as *n (%)*. Neurological outcomes are presented hierarchically, with overall brain lesion status reported as the primary outcome and lesion subtypes shown descriptively.

3.2 Nutritional management

Nutritional management followed a standardized departmental protocol and was applied consistently across the study cohort. Immediately after birth, parenteral nutrition was initiated with amino acids and 10% glucose. At approximately 24 hours of life, a multicomponent lipid emulsion (SMOFlipid) comprising soybean oil, medium-chain triglycerides (MCT), olive oil, and fish oil rich in omega-3 fatty acids, 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. Once enteral intake reached 80 mL/kg/day, milk was fortified with Aptamil BMF to provide hydrolyzed bovine proteins and long-chain polyunsaturated fatty acids (LCPUFAs), including docosahexaenoic acid (DHA) and arachidonic acid (AA).

Parenteral lipid administration was discontinued once enteral volumes reached 90-100 mL/kg/day, while amino acids and 10% glucose infusion were continued.

3.3 Plasma sample collection

Venous or capillary blood samples (0.5-1 mL) were collected longitudinally at five predefined timepoints, strictly 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. The second (T1) and third (T2) samples were collected at 48-72 hours of life and at 7 days of life (± 1 day), respectively. When available, the umbilical venous catheter was used for blood withdrawal; otherwise, peripheral venipuncture was performed. Later timepoints included T3, collected at approximately 33 weeks postmenstrual age, and T4, collected at term-equivalent age (around 40 weeks postmenstrual age), both obtained via peripheral venipuncture.

Whole blood was collected into EDTA tubes to prevent coagulation and processed for plasma isolation as soon as possible. To preserve molecular integrity and minimize pre-analytical variability, blood samples were either processed immediately

or temporarily stored at 4 °C for a maximum of 3 hours prior to centrifugation. Plasma separation was performed by centrifugation for 15 minutes at 1200 × g using a benchtop centrifuge. After centrifugation, plasma aliquots were anonymized using alphanumeric sample identifiers and stored at −80 °C until analysis.

3.4 Automated extracellular vesicle enrichment for integrated proteomic and lipidomic sample preparation

EVs were enriched from plasma samples using a semi-automated magnetic bead-based workflow. The protocol was adapted from the Mag-Net method – originally designed for plasma proteomics[14] – to enable the simultaneous extraction of lipids and proteins from minimal sample volumes. All steps were performed on a KingFisher™ Apex magnetic handling station (Thermo Fisher Scientific) at 4 °C to ensure the biochemical stability of the EV cargo.

For each sample, 10 µL of plasma were thawed on ice and diluted to 100 µL using 40 µL of PBS followed by a 1:1 dilution with Bis-Tris Propane (BTP) binding buffer (100 mM BTP, pH 6.3, 150 mM NaCl) to promote electrostatic interactions between negatively charged EV membranes and strong anion exchange (SAX) magnetic beads. 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).
Following sequential washes to remove non-vesicular plasma components, the bead-bound EVs were transferred to an extraction plate containing 600 µL 1-butanol (BuOH)/acetonitrile (ACN)/water (H_2O) (3:1:1, v/v)[15] as monophasic extraction solvent. This step promotes membrane disruption and the subsequent release of the molecular cargo. Samples were incubated overnight at −20 °C to induce protein precipitation onto the magnetic beads while maintaining lipid species in the soluble organic phase.

On the following day, samples were reintroduced into the KingFisher™ Apex workflow for a second processing round. 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, 25 min at 4,000 × g), transferred to clean tubes, vacuum-dried, and stored at −80 °C until LC-MS/MS analysis.

In parallel, the bead-bound proteins underwent automated processing, including sequential washes with ACN, 70% ethanol, and isopropanol (IPA) to remove contaminants. On-bead digestion was performed in 100 mM Tris-HCl (pH 8.0) supplemented with 10 mM tris(2-carboxyethyl)phosphine (TCEP) and 4 mM chloroacetamide (CAA) for reduction and alkylation. Proteolysis was carried out 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. The resulting peptide mixtures were acidified with 10 µL of 2% (v/v) trifluoroacetic acid (TFA) and stored at −20 °C until LC-MS/MS analysis.

3.5 Data acquisition

3.5.1 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 a BuOH/IPA/H_2O mixture (8:23:69, v/v), supplemented with 5 mM phosphoric acid, and 5 µL of each sample were directly injected onto a reversed-phase column.

Chromatographic separation was achieved on an ACQUITY UPLC BEH C18 column (1.7 µm, 2.1 × 50 mm; Waters), maintained at 65 °C. Lipids were separated using a reversed-phase gradient with a total run time of 7 minutes 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), while 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) in positive ionization mode, using a spray voltage of +3.5

kV. The inlet capillary temperature was set to 300 °C. Nitrogen was used as sheath gas 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 a resolution of 70,000 (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 a resolution of 17,500, 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.

3.5.2 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. For each sample, 20 µL of peptide digest were directly loaded onto EvoTips according to the manufacturer's instructions. Briefly, dry tips were activated with solvent B (0.1% FA in ACN), conditioned with IPA, and equilibrated with solvent A (0.1% FA in H_2O). Peptide samples were then loaded onto the tips, washed with solvent A, and kept wet until analysis.

Peptide separation was achieved on an Aurora Rapid 8 cm x 75 µm UHPLC column (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.

Mass spectrometry data were acquired in data-independent acquisition (DIA) mode. The FAIMS compensation voltage (CV) was set to −45 V at standard resolution. Full MS scans were acquired at a resolution of 60,000 over a mass range of 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 of 15 Da width with a 1 Da overlap. Fragment ion spectra were acquired at a resolution of 15,000, with the injection time set to Auto and a normalized collision energy of 25%. Raw data were acquired in profile mode using positive ion polarity.

3.6 Data processing and quality control

3.6.1 Lipidomic data processing

Raw lipidomics data were processed using MS-DIAL v4.8 for peak detection, deconvolution, retention time alignment, and lipid annotation[16]. The resulting master area table was exported as a tab-delimited text file and imported into MS-FLO v1.8 to perform additional deisotoping and to resolve redundant or duplicated lipid features, thereby improving annotation consistency[17].

The refined data table was imported into R v4.5.2 for downstream data cleaning, quality control, normalization, and missing value handling using custom scripts.
Putative contaminants and non-biological signals were removed prior to statistical analysis. Lipid features annotated as chemical reagents, plasticizers, polymer-related compounds, or background contaminants were excluded based on annotation keywords and pattern-based filtering.
Quality control (QC) filtering was performed using pooled QC samples. For each lipid feature, the mean intensity and coefficient of variation (CV) across all QC injections were calculated. Features showing poor technical reproducibility were removed by applying a CV threshold of 30%.
To retain biologically informative lipid species, a fill-rate-based filtering strategy was applied. For each lipid feature, the proportion of non-missing values was calculated separately within each time point. Features were retained if they exhibited a minimum fill rate of 70% (≥70% non-missing values) in at least one time point.
When multiple features were annotated with the same lipid name, duplicate features were resolved by retaining a single representative feature per lipid species. Selection was based on a hierarchical criterion prioritizing: (i) highest maximum fill rate across time points, (ii) lowest QC CV, and (iii) highest mean QC intensity.
Following feature-level filtering, raw intensities were log2-transformed to stabilize variance and approximate normality.
To correct for sample-wise intensity shifts due to technical variability (e.g. injection or loading effects), median-centering normalization was applied. For each sample, the median log2 intensity was aligned to the global median calculated across all

samples, under the assumption that the majority of lipid species do not undergo systematic global changes.
Missing values were subsequently imputed assuming left-censored missingness, consistent with low-abundance lipid species falling below the detection limit in LC-MS experiments. For each sample, missing values were replaced with random values drawn from a Gaussian distribution defined by the observed data distribution of that sample, with the mean shifted 1.8 standard deviations towards lower intensities and the standard deviation scaled to 30% of the original variance.

The resulting normalized and imputed lipidomics dataset was used for all downstream statistical and integrative analyses.

3.6.2 Proteomic data processing

Raw proteomics data were processed using Spectronaut v18 (Biognosys AG) in directDIA mode with default settings[18]. Enzymes/Cleavage Rules were set to Trypsin/P, LysC. Peptide and protein identification was performed by searching against the Uniprot human reference proteome, using carbamidomethylation of cysteine residues as a fixed modification, while methionine oxidation and protein N-terminal acetylation were selected as variable modifications. A false discovery rate (FDR) threshold of 1% was applied at the 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. Protein quantification was performed at the MS2 level, using Spectronaut default settings for directDIA-based label-free quantification. The resulting protein-level pivot table was exported for downstream processing.

Protein-level data were imported into R v4.5.2 for data curation, quality control, normalization, and missing value handling using custom scripts. Protein identifiers were harmonized by extracting a single representative UniProt accession and gene symbol per protein group, prioritizing reviewed Swiss-Prot entries when available. Sample identifiers were standardized and matched to biological metadata to ensure consistent alignment between expression data and experimental variables.

To reduce sparsity while preserving biologically informative proteins, a fill-rate–based filtering strategy was applied. Proteins were retained if they showed at least 70% non-missing values in at least one time point.

Protein intensities, already reported as log2-transformed values by Spectronaut, were normalized by median centering to correct for sample-wise technical variability. For each sample, the median log2 intensity was aligned to the global median across all samples, under the assumption that the majority of proteins do not undergo systematic global abundance changes.

Missing values were handled using a hybrid imputation strategy accounting for different missingness mechanisms. Values classified as missing at random (MAR) were imputed using structured least squares adaptive (SLSA) imputation, incorporating time point information. Conversely, values classified as missing not at random (MNAR), consistent with low-abundance proteins falling below the detection limit, were imputed using a downshifted Gaussian distribution.

The resulting curated proteomics dataset was used for all downstream statistical analyses and multi-omics integration.

3.7 Data analysis

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

Longitudinal molecular changes were assessed using linear mixed-effects models (LMMs). For each protein or lipid, abundance was modeled as a function of sampling time point (T0-T4), with gestational age at birth included as a fixed-effect covariate. Infant identity was included as a random intercept to account for repeated measurements. A global test was used to assess the effect of time, and p-values were corrected using the Benjamini-Hochberg (BH) procedure. Features with FDR_global < 0.05 were considered significantly time-regulated. For endpoint analyses, the contrast between term-equivalent age (T4) and birth (T0) was

extracted from each model. Effect sizes were expressed as β coefficients. Features were considered significantly regulated at T4 versus T0 if they met FDR < 0.05 and |β| > 0.58.

Cross-omics integration was performed using only temporally regulated features (FDR_global < 0.05). Protein and lipid abundance profiles were z-score standardized, and pairwise Spearman correlations were computed across paired longitudinal samples. Correlation p-values were corrected using BH adjustment.
The full correlation landscape was retained for descriptive analyses. Network edges were defined by applying FDR < 0.05 and |ρ| ≥ 0.35. The resulting weighted, undirected network was subjected to 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 time point 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. Proteins were selected based on the longitudinal LMM results and the T4 versus T0 contrast, using FDR < 0.05 and an effect size threshold (|β| > 0.58). Enrichment analyses were conducted using all quantified proteins as background, with multiple testing correction applied using the BH method (FDR < 0.05). Lipid ontology enrichment was performed using the LION web-based tool, ranking lipid species by their T4 versus T0 β coefficients and retaining significantly enriched terms (FDR < 0.05).

4. Results

4.1 Overview of the longitudinal plasma EV multi-omics dataset

To investigate how the molecular composition of plasma-derived extracellular vesicles (EVs) changes from very preterm birth to term-equivalent age (TEA), we analyzed longitudinal plasma EV samples from a cohort of 16 very preterm infants born before 32 weeks of gestational age.

Plasma samples were collected at up to five predefined time points: at birth (T0), at 48-72 hours of life (T1), at 7 days of life (T2), at approximately 33 weeks postmenstrual age (T3), and at TEA, corresponding to 40 weeks postmenstrual age (T4). Only infants with samples available for at least four time points were included. Demographic and clinical characteristics of the study population are reported in **Table 1**. An overview of the longitudinal study design, sampling timeline, and experimental workflow is schematically illustrated in **Fig.1a**.

At each time point, EVs isolated from the same plasma sample were analyzed in parallel by proteomics and lipidomics, allowing protein and lipid changes to be followed over time within the same infants. After quality filtering, a total of 1,528 EV-associated protein groups and 421 unique lipid species were consistently identified and quantified across the longitudinal dataset.

Proteomic profiling revealed a wide dynamic range of EV-associated proteins, with a limited number of highly abundant components and a large set of lower-abundance proteins (**Fig. 1b**). Subcellular localization enrichment analysis showed a strong over-representation of extracellular and vesicle-related compartments, consistent with the EV origin of the analyzed material (**Fig. 1c**). In parallel, lipidomic profiling revealed a membrane-enriched composition dominated by glycerophospholipids and

sphingolipids, with sphingomyelins representing the most abundant lipid class (**Fig. 1d**).

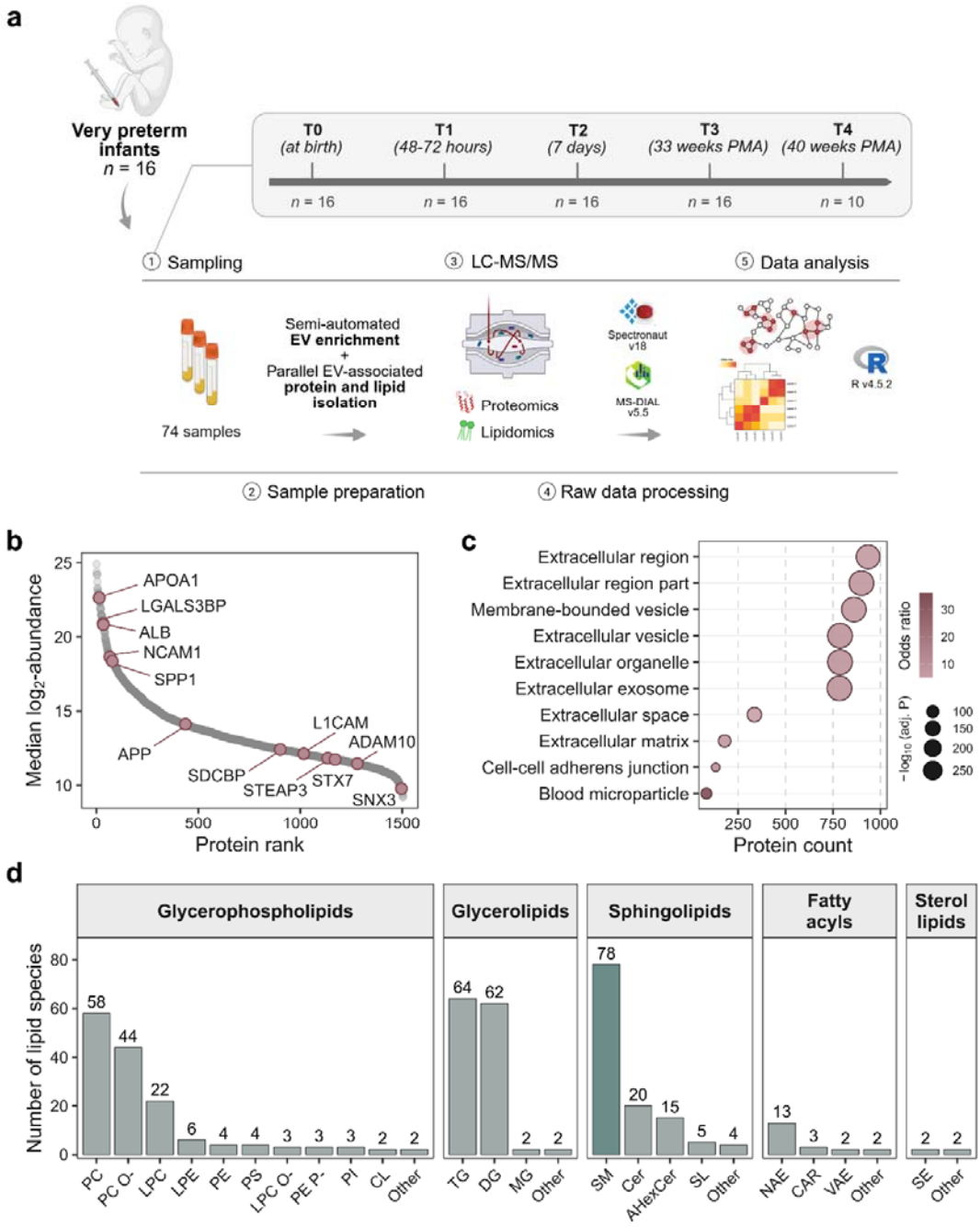

Fig. 1: Longitudinal study design and plasma EV multi-omics profiling in very preterm infants. (a) Overview of the longitudinal cohort, sampling timeline, and analytical workflow. Plasma samples were collected from very preterm infants born at <32 weeks of gestational age (n = 16) at five time points from birth to term-equivalent age: at birth (T0), 48-72 hours of life (T1), 7 days of life (T2), 33 weeks postmenstrual age (PMA, T3), and 40 weeks PMA (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 analyses were performed by LC-MS/MS, with raw data processed using dedicated pipelines prior to downstream statistical analyses. **(b)** Dynamic range of the plasma EV proteome, shown as a rank-abundance plot of median log protein abundance across all samples, with selected EV-associated proteins highlighted. **(c)** Subcellular localization enrichment of the EV proteome based on the Jensen COMPARTMENTS database. Terms are ordered by the number of associated proteins along the x axis (protein count). Dot size reflects statistical significance ($-\log_{10}$ FDR), while color intensity indicates odds ratio. **(d)** Distribution of identified lipid species across major lipid classes, grouped into glycerophospholipids, glycerolipids, sphingolipids, fatty acyls, and sterol lipids.

4.2 Plasma EV proteomic and lipidomic trajectories from very preterm birth to term-equivalent age

To characterize changes in EV molecular profiles across postnatal development, we analyzed both the proteomic and lipidomic datasets using a linear mixed-effects modeling (LMM) framework. This approach was chosen to account for the longitudinal study design, with repeated measurements collected from the same individuals across multiple postnatal time points. Infant-specific random intercepts were included to model within-subject correlation across time. Sampling time point (T0-T4) was included as a fixed effect to capture systematic developmental trends. 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 first tested for a global effect of time across all sampling points and subsequently focused on the contrast between term-equivalent age (TEA, T4) and birth (T0) as the primary developmental comparison. Multiple testing was controlled using the Benjamini-Hochberg procedure.

Extensive postnatal remodeling of the EV proteome

Applying this framework to the plasma EV proteome revealed extensive temporal remodeling following very preterm birth. Overall, 875 of the 1,528 quantified proteins showed a significant global effect of time (FDR < 0.05), indicating that a substantial fraction of EV cargo is dynamically regulated during postnatal development.

Unsupervised hierarchical clustering of time-regulated proteins identified four main clusters that could be grouped into two opposing trajectories: proteins that progressively decreased from birth to TEA (clusters 1A and 1B) and proteins that increased over time (clusters 2A and 2B) (**Fig. 2a,b**). These trajectories indicate that EV proteomic changes are coherent and directional, rather than reflecting unspecific longitudinal variability.

Focusing on the net developmental shift between TEA and birth, we identified 439 proteins significantly downregulated and 127 proteins significantly upregulated at TEA (FDR < 0.05, effect size |β| > 0.58; **Fig. 2c**). Proteins that decreased toward TEA were mainly associated with RNA metabolism, translation, and early growth-related processes, including core components of the translational machinery such as RPS6 and EIF4A2, as well as developmentally regulated proteins such as AFP (**Fig. 2d,e**). In contrast, proteins increasing toward TEA were significantly enriched for immune-related functions, including complement activation and immunoglobulin-mediated responses (**Fig. 2e**). Representative examples include complement-associated proteins such as C9 and PROS1, which progressively increased across postnatal development, consistent with the emergence of a more mature systemic immune-related EV profile (**Fig. 2d**).

Selective structural and metabolic remodeling of the EV lipidome

We next examined longitudinal changes in the plasma EV lipidome using the same LMM strategy. Compared with the proteome, lipidomic remodeling involved a smaller subset of features but remained clearly structured over time. Out of 421 quantified lipid species, 100 showed a significant global effect of time (FDR < 0.05; **Fig. 3a**).
Unsupervised hierarchical clustering of time-regulated lipids resolved four distinct longitudinal trajectories, each comprising lipid species from multiple chemical classes (**Fig. 3b**). This organization indicates that EV lipidomic changes occur through coordinated programs affecting defined groups of molecular species, rather than uniform shifts across entire lipid classes.

Comparison of TEA with birth identified 41 lipid species significantly upregulated and 22 significantly downregulated at TEA (FDR < 0.05, effect size |β| > 0.58; **Fig. 4c**). Lipids increasing over development were dominated by triacylglycerols (TG), ether-

linked phosphatidylcholines (PC O-) and selected sphingomyelins (SM). Among the strongest increases were SM 28:1;O2 and ether-linked PCs such as PC O-39:8 and PC O-37:8, together with multiple TG species, consistent with progressive enrichment of neutral lipid cargo alongside selective remodeling of EV membrane components.

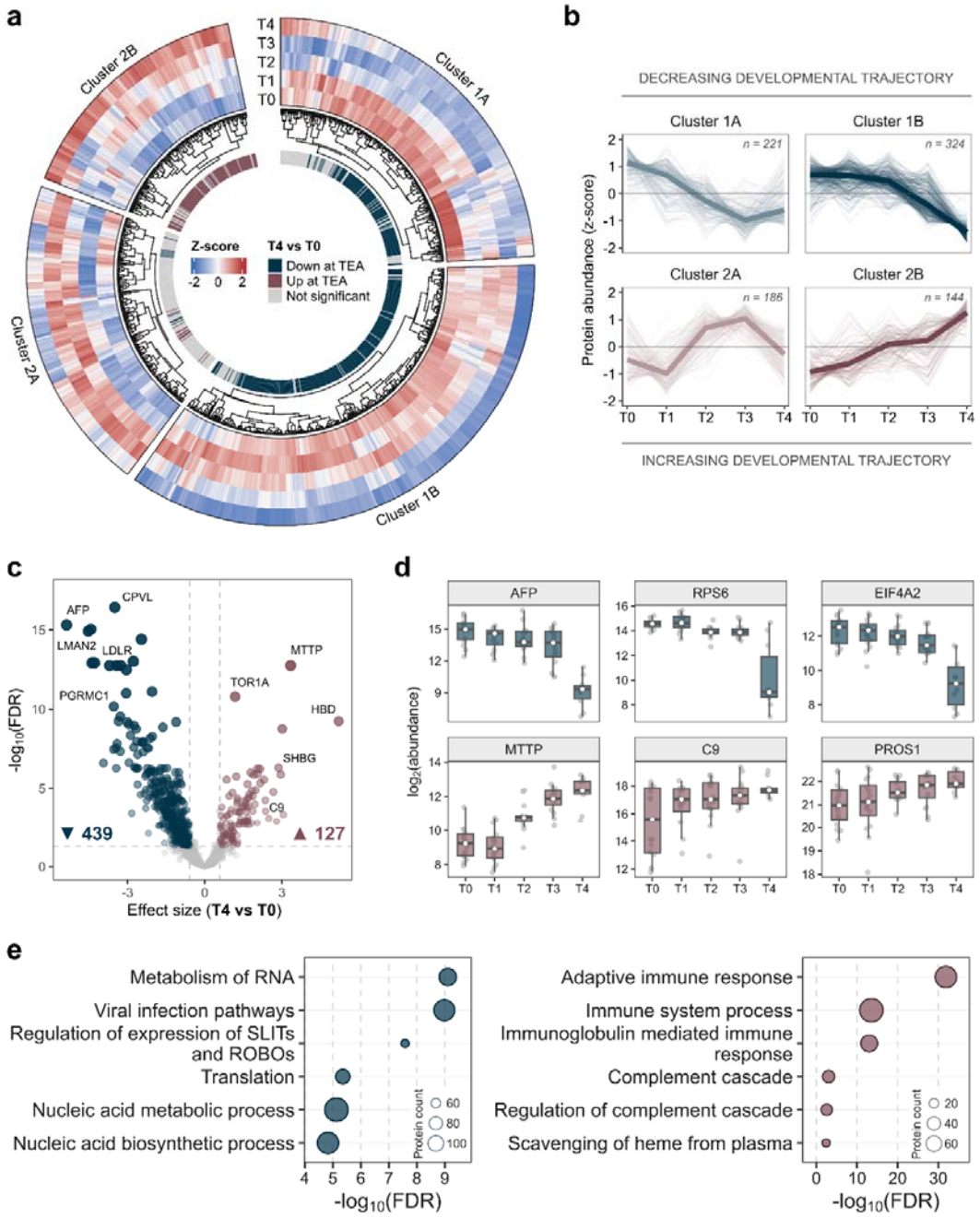

Fig. 2: Longitudinal remodeling of the plasma EV proteome from very preterm birth to term-equivalent age. (a) Heatmap of EV proteins showing a significant temporal effect in the longitudinal linear mixed-effects model (LMM; FDR < 0.05). Protein abundances are z-scored by row across five

time points, from birth (T0) to term-equivalent age (TEA, T4), and organized by hierarchical clustering. The outer ring indicates the direction of change at TEA relative to birth (T4 vs T0). **(b)** Cluster-specific developmental trajectories of protein abundance across time points. Thin lines represent individual proteins and bold lines indicate cluster medians, highlighting coordinated decreasing (clusters 1A and 1B) and increasing (clusters 2A and 2B) temporal programs. **(c)** Volcano plot of the T4 (TEA) vs T0 contrast derived from the LMM. Each point represents a protein; the x axis shows effect size (β coefficient) and the y axis statistical significance ($-\log_{10}$ FDR). Proteins significantly decreased or increased at TEA are highlighted. Numbers indicate the total count of significantly downregulated (439) and upregulated (127) proteins (FDR < 0.05, $|\beta| > 0.58$). **(d)** Representative proteins illustrating distinct longitudinal patterns. Boxes represent median and interquartile range; whiskers indicate data range. **(e)** Functional enrichment analysis of downregulated (left) and upregulated proteins (right) at TEA vs birth. For each group, the top enriched terms of Gene Ontology Biological Process and Reactome databases are shown, ordered by statistical significance along the x axis ($-\log_{10}$ FDR); dot size reflects the number of proteins associated with each term.

Conversely, lipids reduced toward TEA were enriched for diacyl PCs, phosphatidylserines (PS), diacylglycerols (DG), and long-chain SM. 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. Within individual lipid classes, regulation was highly selective rather than uniform. In particular, SM 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, lipid ontology (LION) enrichment analysis was performed on lipids ranked by their T4 versus T0 effect sizes (**Fig. 3d**). Lipids enriched at T4 were associated with annotations related to lipid storage and lipid droplet biology, whereas lipids reduced over time were preferentially associated with membrane structural properties, including bilayer thickness and curvature.

Finally, to assess whether these changes reflected specific structural trends, significantly regulated lipids from major EV membrane classes – PC, PC O-, and SM – were mapped onto carbon number-double bond space (**Fig. 3e**). 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 selective redistribution among molecular species.

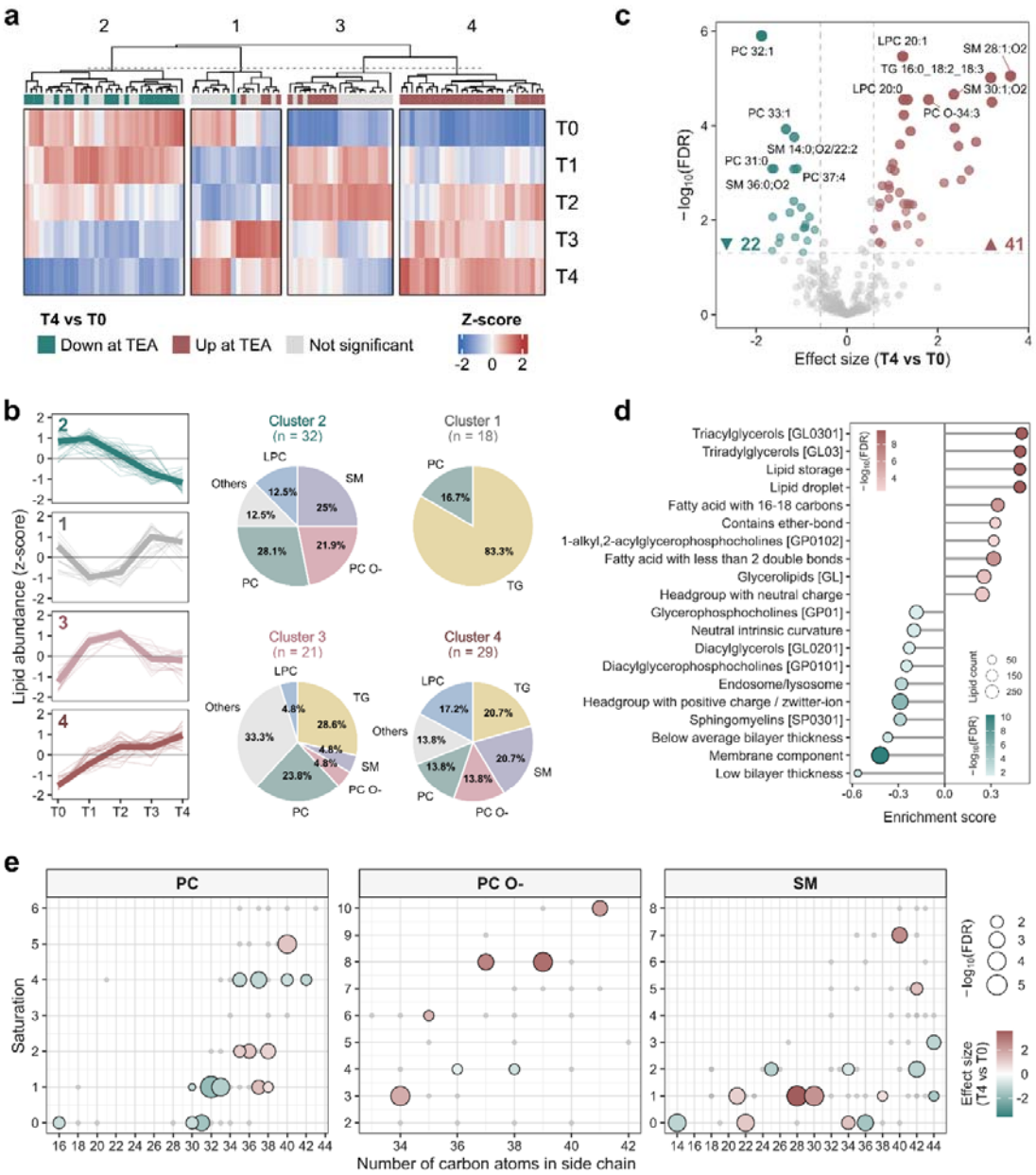

Fig. 3: Longitudinal remodeling of the plasma EV lipidome from very preterm birth to term-equivalent age. (a) Heatmap of lipid species showing a significant temporal effect in the longitudinal linear mixed-effects model (LMM; FDR < 0.05). Lipids are z-scored by row and grouped by unsupervised hierarchical clustering into four longitudinal trajectory patterns. The right-side annotation indicates the direction of change at term-equivalent age (TEA, T4) relative to birth (T0). **(b)** Cluster-specific developmental trajectories of EV lipid abundance across time points. Thin lines represent

individual lipid species, while bold lines indicate cluster medians. The lipid class composition of each cluster is shown on the right as pie charts. **(c)** Volcano plot of the T4 (TEA) vs T0 contrast derived from the LMM. Each point represents a lipid species; the x axis shows effect size (β coefficient) and the y axis statistical significance ($-\log_{10}$ FDR). Lipids significantly increased or decreased at TEA are highlighted. Numbers indicate the total count of significantly upregulated (41) and downregulated (22) lipids (FDR < 0.05, |β| > 0.58). **(d)** Lipid ontology (LION) enrichment analysis based on the β coefficient (T4 vs T0). Enrichment scores are shown as lollipop plots. Dot size indicates the number of annotated lipid species per term, and dot color intensity represents statistical significance ($-\log_{10}$ FDR), with separate color scales for lipids increased or decreased at TEA. **(e)** Acyl-chain structural space plot showing the distribution of significantly regulated phosphatidylcholines (PC), ether-linked phosphatidylcholines (PC O−), and sphingomyelins (SM) across carbon number-double bond space. Each dot represents one lipid species; dot size reflects statistical significance ($-\log_{10}$ FDR) and color encodes effect size (T4 vs T0), highlighting class-specific structural redistribution across postnatal development.

Together, these analyses indicate that postnatal development in very preterm infants is associated with progressive remodeling of plasma EV composition across both omics layers. Proteomic changes are widespread and involve coordinated shifts in biological pathways related to growth and immune maturation, whereas lipidomic changes are more selective and reflect redistribution among specific lipid species. Overall, plasma EVs capture developmental molecular changes occurring during the very preterm period.

4.3 Network-based integration reveals coordinated proteomic and lipidomic EV modules

To identify coordinated molecular programs across the EV proteome and lipidome, the two omics layers were integrated using a cross-omics correlation network approach (**Fig. 5a**). The network was constructed starting from features showing a significant global effect of time in the longitudinal LMM analysis (875 proteins and 100 lipids in total), thereby focusing on the dynamically regulated component of EV remodeling during postnatal development.

Inter-omic coordination was assessed by computing pairwise Spearman correlations between proteins and lipids across paired longitudinal samples, explicitly restricting the analysis to cross-layer associations. This approach captures similarity in

temporal trajectories between molecular layers rather than correlations driven by shared absolute abundance or global time trends. After multiple-testing correction and effect-size filtering, the resulting network consisted of protein and lipid nodes connected by statistically significant positive or negative correlations (**Fig. 5b**). The bimodal distribution of retained correlations, with a clear separation between positive and negative associations, indicates structured and non-random inter-omic coordination during EV maturation.

Louvain community detection was applied to the filtered network, identifying five distinct cross-omics modules (M1–M5) composed of both proteins and lipids (**Fig. 5a,c**). Each module represents a set of molecular features exhibiting coherent longitudinal behavior across omics layers. While proteins accounted for the majority of nodes in all modules, lipids were consistently integrated within each module rather than forming isolated clusters, highlighting tight coupling between EV protein and lipid remodeling.

Module size and composition varied substantially, ranging from 115 to 237 nodes (**Fig. 5c**). Importantly, lipids within modules were not randomly distributed by class but reflected the selective lipidomic remodeling observed at the single-feature level, with specific modules enriched for ether-linked phosphatidylcholines, sphingomyelins, or triacylglycerols.

Functional annotation of protein components revealed distinct biological themes across modules (**Fig. 5d**). Module M1 was enriched for processes related to protein turnover and intracellular organization, consistent with the progressive decline of early growth-associated EV cargo. Module M2 was characterized by immune system activation and complement-mediated responses, reflecting maturation of innate immune signaling. Module M3 showed enrichment for adaptive immune functions, whereas Module M4 was associated with translation- and ribosome-related processes. Module M5 was enriched for extracellular matrix organization and tissue remodeling, suggesting a link between EV molecular remodeling and developmental tissue maturation.

Together, these findings demonstrate that longitudinal EV proteomic and lipidomic changes are organized into coordinated cross-omics modules rather than occurring independently within individual molecular layers. This network-based integration

reveals structured molecular programs underlying EV maturation in very preterm infants and provides a biologically interpretable framework for linking EV molecular trajectories to clinical phenotypes.

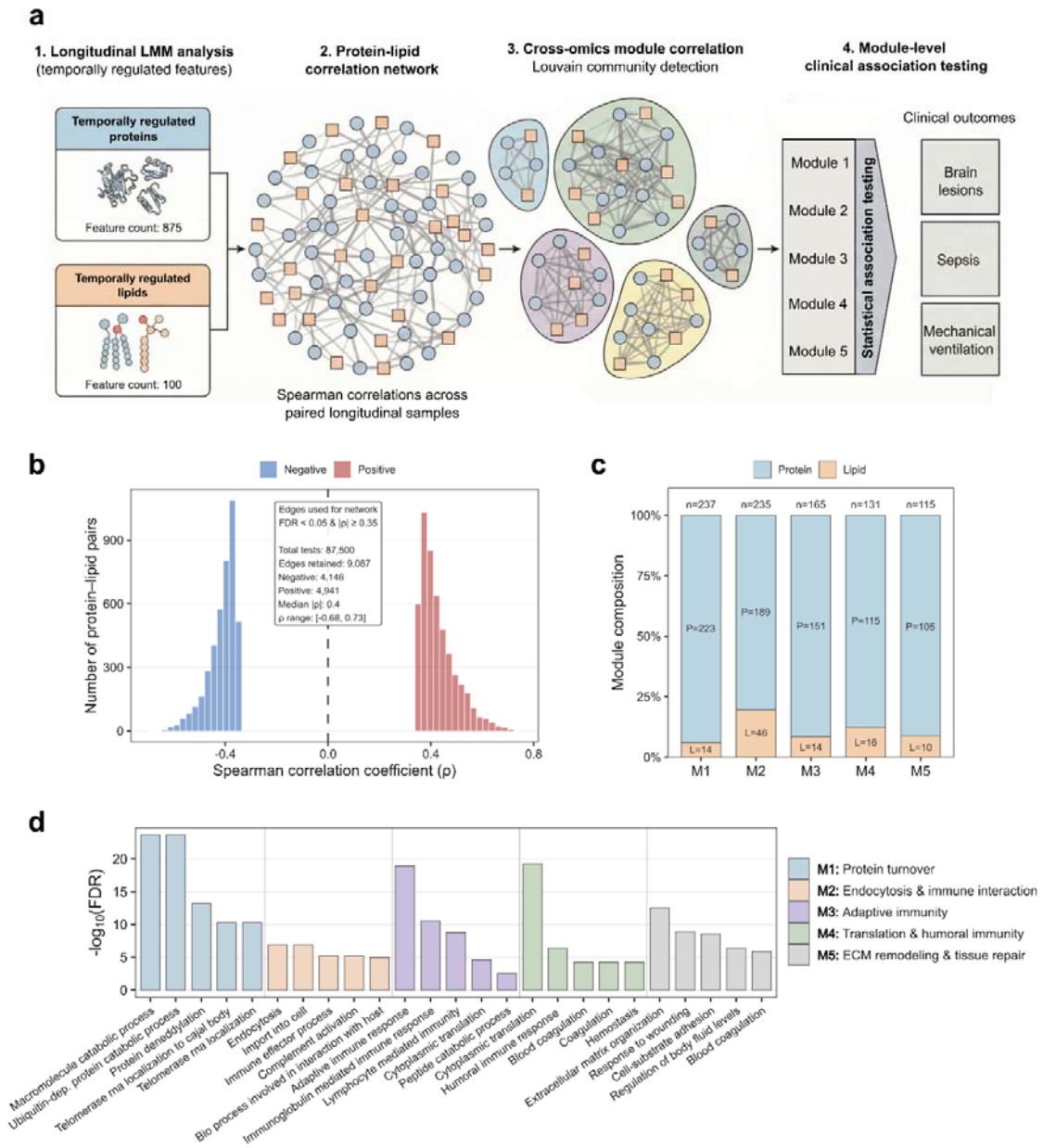

Fig. 4: Cross-omics network integration identifies coordinated EV proteomic-lipidomic modules.

(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 for protein-lipid pairs retained in the network after multiple-testing correction and effect-size filtering ($|\rho|>0.35$). Positive and negative correlations are shown separately, highlighting a structured and bimodal distribution of inter-omic associations. **(c)** Protein and lipid composition of the five cross-omics modules (M1-M5). Bars indicate the relative contribution of proteins and lipids within each module, with total node counts reported above each bar. **(d)** Functional annotation of protein components within each module based on Gene Ontology Biological Process enrichment analysis. Bars represent $-\log_{10}$(FDR) values for selected enriched terms, highlighting distinct biological themes across modules, including protein turnover, immune-related processes, translation, and extracellular matrix remodeling.

4.4 Association of EV molecular modules with clinical outcomes

Associations between EV molecular modules and clinical outcomes were explored using a module-level CAMERA approach, which tests whether groups of correlated molecular features show a coordinated association with an outcome while accounting for inter-feature correlation. This strategy was chosen to increase robustness compared with single-feature testing in this exploratory cohort.

At the protein level, significant module-level associations were detected for brain injury-related outcomes (**Fig. 5**). Module M1 showed a strong negative association with hemorrhagic lesions (FDR = $1.6 \times 10^{-}$□), indicating a coordinated decrease of this module in infants with hemorrhagic brain injury. Proteins contributing to this signal included complement regulatory components such as C4BPA and C4BPB, as well as hemoglobin subunit HBM, suggesting alterations in EV-associated complement regulation and blood-derived components in this condition.

Module M3 displayed consistent positive associations with both hemorrhagic lesions (FDR = $4.5 \times 10^{-}$□) and the broader outcome of any brain lesion, encompassing hemorrhagic lesions and white matter injury (FDR = $3.7 \times 10^{-}$□). This module is enriched for immune-related proteins, and the association was driven by coordinated increases in multiple immunoglobulin components (IGHV, IGKV, IGLV families), the joining chain JCHAIN, the proteasome subunit PSMB8, and immune-associated proteins such as GCA, GLIPR2, alongside a decrease in SELL. Together, these features indicate a coherent EV immune signature associated with brain injury, rather than isolated protein-level effects.

At the lipid level, no individual lipid species reached nominal significance. However, module-level analysis revealed coordinated associations that were not detectable at the single-feature level. Module M2 showed a significant negative association with hemorrhagic lesions (FDR = 0.020), reflecting a concerted decrease of lipids within this module. Although individual effects were modest, contributing lipids included ether-linked phosphatidylcholines (PC O-37:8, PC O-39:8), lysophosphatidylcholines, and selected sphingomyelins, consistent with subtle but coordinated alterations in EV membrane-related lipid composition.

In addition, lipid module M3 was positively associated with sepsis (FDR = 0.040) and with mechanical ventilation (FDR = 0.012). This module comprised a small set of lipids, including phosphatidylcholines, diacylglycerols, sphingomyelins, and triacylglycerols, none of which showed strong individual associations. The module-level signal therefore reflects weak but consistent directional changes across multiple lipid species, suggesting a diffuse EV lipid response in these clinical contexts.

Overall, these analyses indicate that associations between EV molecular profiles and clinical outcomes in very preterm infants are more robustly captured at the module level than at the level of individual proteins or lipids. While protein modules showed clearer and more consistent associations with brain injury-related outcomes, lipid modules revealed subtler coordinated patterns linked to hemorrhagic lesions, sepsis, and mechanical ventilation. These findings support the use of module-based approaches to detect distributed molecular signals in exploratory EV multi-omics studies.

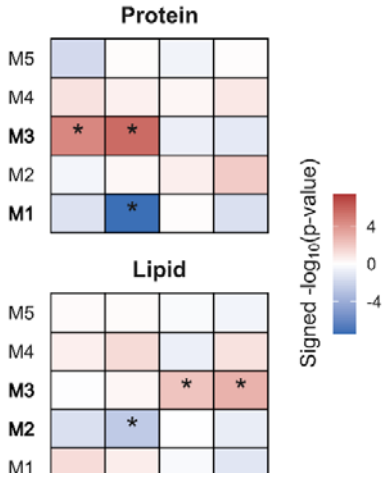

Fig. 5: Integrated EV modules show significant associations with clinical phenotypes in preterm infants. Heatmap summarizing module-level associations between EV molecular modules and clinical outcomes, assessed using a CAMERA-based competitive gene set framework. Protein (top) and lipid (bottom) groups 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 individual lipid species did not reach significance, coordinated shifts across lipid modules were detected at the module level, highlighting the ability of module-based approaches to capture subtle but consistent EV molecular patterns.

5. Discussion and conclusions

This study shows that EVs provide a focused and biologically meaningful window into postnatal adaptation in very preterm infants. By analyzing EV proteins and lipids in parallel and in a longitudinal framework, we captured coordinated molecular programs that reflect both developmental maturation and clinical stress during early life.

A central observation is that EV remodeling is structured rather than diffuse. Proteomic changes indicate a transition from an early growth- and translation-dominated profile toward a progressively more immune-competent state. This shift mirrors known aspects of extra-uterine adaptation in very preterm infants and suggests that EVs reflect systemic maturation. Importantly, these signals emerge at the level of EV cargo, a compartment enriched for cell-derived and functionally relevant molecules, rather than being diluted within whole plasma.
EV lipid remodeling followed a different but complementary logic. Changes were selective and strongly structure-dependent, affecting specific lipid species rather than entire classes. The enrichment of triacylglycerols and ether-linked phosphatidylcholines toward TEA suggests evolving energetic and membrane requirements, while the redistribution of membrane-associated lipids points to refinement of EV structural properties. These patterns argue against non-specific metabolic effects and instead support regulated lipid packaging during EV maturation.

Integrating proteomic and lipidomic data was essential to uncover these dynamics. Cross-omics analysis showed that proteins and lipids co-vary in coordinated modules with shared temporal behavior. These modules summarize complex longitudinal data into biologically interpretable units and reveal molecular programs that are not evident when each omics layer is analyzed separately. From a clinical

perspective, this is particularly relevant in small and heterogeneous neonatal cohorts, where single-feature signals are often weak but coordinated changes across pathways are robust.

The association between EV molecular modules and clinical features such as brain injury, sepsis, and respiratory support further supports the clinical relevance of EV profiling. While causal inference is beyond the scope of this study, these findings suggest that EV composition reflects integrated responses to developmental stress and injury. EVs therefore emerge as promising carriers of clinically meaningful molecular information in fragile neonatal populations.

A key strength of this approach is its compatibility with neonatal constraints. EV profiling enables longitudinal monitoring from minimal plasma volumes and captures system-level information derived from multiple tissues. Compared with whole plasma, EVs offer improved signal specificity by concentrating cell-derived molecular cargo and reducing dominance of highly abundant circulating components. This makes EV-based multi-omics particularly suited for studying complex biology in preterm infants.

However, several limitations should be considered. The cohort size is necessarily limited, reflecting the challenges of longitudinal sampling in very preterm infants, and larger studies will be required to validate these findings and assess their generalizability. Sampling was unbalanced, with fewer infants at term-equivalent age (T4), which may reduce statistical power for late-timepoint effects and outcome associations; however, the use of longitudinal mixed-effects models partially mitigates this limitation by leveraging within-subject trajectories across earlier time points.

EV heterogeneity was not resolved at the subtype or cell-of-origin level. While this limits mechanistic attribution, it is consistent with the study's primary aim of capturing EVs as integrated systemic readouts rather than cell-specific signals. Finally, the study is exploratory and descriptive in nature; functional validation will therefore be an important next step to establish causal links between EV remodeling and clinical outcomes.

Future work should focus on extending this framework to larger and multicenter cohorts, and on linking early EV trajectories to long-term neurodevelopmental outcomes.

In conclusion, this study demonstrates that extracellular vesicles represent a powerful alternative to whole plasma for molecular profiling in very preterm infants. Parallel proteomic and lipidomic analysis of EVs captures coordinated, cell-derived molecular programs that reflect both development and disease. EV-based multi-omics therefore provides a strong and scalable foundation for future clinical and translational research in neonatal medicine.

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. Goedicke-Fritz, S. *et al.* Preterm Birth Affects the Risk of Developing Immune-Mediated Diseases. *Front. Immunol.* **8**, 1266 (2017).
5. Shim, S.-Y. & Kim, H.-S. Oxidative stress and the antioxidant enzyme system in the developing brain. *Korean J. Pediatr.* **56**, 107 (2013).
6. Galinsky, R. *et al.* Complex interactions between hypoxia-ischemia and inflammation in preterm brain injury. *Dev. Med. Child Neurol.* **60**, 126–133 (2018).
7. Panfoli, I. *et al.* Oxidative Stress as a Primary Risk Factor for Brain Damage in Preterm Newborns. *Front. Pediatr.* **6**, 369 (2018).
8. 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).
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. 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.
11. 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).
12. Chitoiu, L., Dobranici, A., Gherghiceanu, M., Dinescu, S. & Costache, M. Multi-Omics Data Integration in Extracellular Vesicle Biology—Utopia or Future Reality? *Int. J. Mol. Sci.* **21**, 8550 (2020).
13. 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).

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. Tsugawa, H. *et al.* MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat. Methods* **12**, 523–526 (2015).
17. 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).
18. 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).