Received: 8 May 2020 | Revised: 31 March 2021 | Accepted: 1 April 2021 | Accepted article online: 8 April 2021

DOI: 10.1002/prca.202000038

Proteomics Clinical Applications

RESEARCH ARTICLE

Protein profiles in plasma: Development from infancy to 5 years of age

Maria Mikus[1,2] | Hans Järnbert-Pettersson[3] | Catharina Johansson[3] | Peter Nilsson[1] | Annika Scheynius[3,4,5] | Johan Alm[3,5]

[1] Department of Protein Science, Division of Affinity Proteomics, SciLifeLab, KTH Royal Institute of Technology, Stockholm, Sweden

[2] Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

[3] Department of Clinical Science and Education, Södersjukhuset, Karolinska Institutet, Stockholm, Sweden

[4] SciLifeLab, Karolinska Institutet, Stockholm, Sweden

[5] Sachs' Children and Youth Hospital, Södersjukhuset, Stockholm, Sweden

Correspondence
Maria Mikus, Institute of Environmental Medicine, Karolinska Institutet, Biomedicum 5B, SE-171 77 Stockholm, Sweden.
Email: maria.mikus@ki.se

Funding information
Swedish Research Council, Grant/Award Numbers: 2012-3011, 2016-1324; Swedish Research Council for Working Life and Social Research, Grant/Award Number: 2006-1630; Regional agreement on medical training and clinical research (ALF) between Stockholm County Council and the Karolinska Institutet and the Karolinska University Hospital, Grant/Award Number: 20180307; Centre for Allergy Research Karolinska Institutet; ChAMP (Centre for Allergy Research Highlights Asthma Markers of Phenotype) consortium funded by the Swedish Foundation for Strategic Research, the Karolinska Institutet, AstraZeneca & Science for Life Laboratory Joint Research Collaboration; Vårdal Foundation; Mjölkdroppen Society, Grant/Award Number: 4-1272/2015; Swedish Asthma and Allergy Association's Research Foundation, Grant/Award Number: 2013-0058; Cancer and Allergy Fund, Grant/Award Number: 2018-0923; Ekhaga Foundation, Grant/Award Number: 2015-85

Abstract

Purpose: Little is known about the longitudinal development of different plasma protein levels during early childhood and particularly in relation to lifestyle factors. This study aimed to monitor the plasma proteome early in life and the influence of different lifestyles.

Experimental Design: A multiplex bead-based immunoassay was used to analyze plasma levels of 97 proteins in 280 blood samples longitudinally collected in children at 6, 12, 24, and 60 months of age living in families with an anthroposophic ($n = 15$), partly anthroposophic ($n = 27$), or non-anthroposophic ($n = 28$) lifestyle.

Results: A total of 68 proteins (70%) showed significantly altered plasma levels between 6 months and 5 years of age. In lifestyle stratified analysis, 59 of 97 (61%) proteins were altered over time within one or more of the three lifestyle groups. Nearly half of these proteins (28 out of 59) changed irrespective of lifestyle. The temporal changes represented four longitudinal trends of the plasma proteins during development, also following stratification of lifestyle.

Conclusions and Clinical Relevance: Our findings contribute to understand the development of the plasma proteome under the influence of lifestyle exposures in early childhood.

KEYWORDS
ALADDIN, childhood, lifestyle, longitudinal

Abbreviations: ALADDIN, Assessment of Lifestyle and Allergic Disease During Infancy; HPA, Human Protein Atlas; SOTA, self-organizing tree algorithm

1 | INTRODUCTION

Environment early in life has a significant impact on future health.[1] Proteins are responsible for key biological functions such as signaling and defense, in an age-related balance influenced by genes, lifestyle, and environment.[2,3] However, little is known about the longitudinal development of different plasma protein levels during early childhood, particularly in relation to environmental factors.[4]

The prospective birth cohort Assessment of Lifestyle and Allergic Disease During Infancy (ALADDIN)[5] offers a unique opportunity to study development of different plasma proteins early in life as well as associations between environmental and lifestyle exposures. In 2004, we initiated this cohort by recruiting families with conventional or anthroposophic lifestyle,[5] with the purpose to identify environmental and lifestyle-related factors during pregnancy and early childhood that may constitute risk of or protection against development of different diseases, particularly allergy. Anthroposophic lifestyle is characterized by, among other things, home delivery, organic diet with live lactobacilli, restrictive use of antibiotics and vaccine, and for the infant avoidance of stressful stimuli.[5–8] The aims of this study were to (i) determine the longitudinal development of different plasma protein levels during early childhood and (ii) compare the influence of different lifestyles.

Statement of clinical relevance

Children develop rapidly during the first years of life with physiological changes. The family lifestyle and living environment influence the child with various exposures that may change over time. How the plasma proteome develops longitudinally during early life and in response to lifestyle exposures has been less studied. Understanding normal physiological changes in relation to pathological processes are important for identification of early disease markers. We here studied the levels of plasma proteins longitudinally during the first 5 years of life in children living in families with an anthroposophic, partly anthroposophic, or non-anthroposophic lifestyle. We found that the majority (70%) of the studied 97 plasma proteins changed over time. In addition, we observed altered levels for 59 of the 97 (61%) proteins within one or more of the three different lifestyle groups. Our findings constitute a contribution to map the development of the plasma proteome under the influence of lifestyle exposures in early childhood.

2 | MATERIALS AND METHODS

2.1 | Subjects

The prospective birth cohort study ALADDIN consists of families with an anthroposophic, partly anthroposophic, or non-anthroposophic lifestyle.[5] A total of 330 families were recruited at anthroposophic and conventional healthcare centers in the Stockholm area, Sweden, between September 2004 and November 2007. Families were enrolled in the study at gestational week 25–38 (median week 30). Four children were excluded because of preterm delivery and two because of miscarriage. The lifestyle groups were classified based on choice of maternal-child health centers and parental responses to a questionnaire 2 months after the birth of the child: The degree of adherence to an anthroposophic lifestyle was categorized as "anthroposophic," "partly anthroposophic," or "non-anthroposophic," based upon the choice of antenatal clinic and parental responses to three questions: (1) "What kind of preschool/school will your newborn child probably go to?" (2) "Has any of the parents, no matter which type of school you have planned for your child, an anthroposophic view of life?," and (3) "Is the family's everyday life influenced by an anthroposophic view of life?" Families answering "anthroposophic school" to question 1 and "yes" to questions 2 and 3 and also attending anthroposophic antenatal clinics were defined as "anthroposophic." Families answering conventional or any other non-anthroposophic type of school to question 1, "no" to questions 2 and 3 and going to conventional antenatal clinic were defined as "non-anthroposophic." Any other combination of answers was defined as "partly anthroposophic."[5] In the present study, we included 70 children from the three different lifestyle groups (15 anthroposophic, 27 partly anthroposophic, and 28 non-anthroposophic), from whom longitudinal plasma samples collected at 6, 12, 24, and 60 months of age were available. Information on diet and other exposures were collected by a questionnaire at 2 months and interviews at 2, 6, 12, 24, and 60 months of age. Data on delivery was retrieved from the Swedish Medical Birth Register, and vaccination data from medical records at child healthcare centers.[7] This study was performed in accordance with the principles expressed in Declaration of Helsinki and approved by the Regional Ethical Review Board in Stockholm (project Dnr 2002-01-07, 474/01 and 2010-04-30, 741-32). All parents gave their written informed consent.

2.2 | Collection of blood samples

Blood samples were collected venously in 4 mL sodium heparin tubes from the children at 6 and 12 months and in 9 mL sodium heparin tubes at 24 and 60 months of age. The blood samples were stored at room temperature before centrifugation for 10 min at 800 × *g* and the plasma was subsequently collected. The distribution of time periods at room temperature across lifestyles were similar, see Table S1. The median volumes of plasma obtained were 1.5 mL (range 0.5–2.5 mL) at both 6 and 12 months, 4.0 mL (range 0.8–4.0 mL) at 24 months, and 3.6 mL (range 2.0–4.0 mL) at 60 months. Plasma was stored at −80°C. Blood was collected venously from all children except from seven children (10%) at 6 months of age and five children (7%) at 12 months of age

where capillary blood was collected due to practical and humanity reasons. For a distribution of capillary blood samples across lifestyles, see Table S2.

2.3 | Protein selection and plasma profiling using antibody suspension bead arrays

For the protein profiling, a panel was designed based on in-house knowledge combined with database searches (Ensembl, Genotator, the Human Protein Atlas [HPA], Ingenuity Pathway Analysis, Intomics, and MetaCore), in general focusing on proteins expressed or detectable in blood/serum/plasma and involved in biological processes of the immune system and inflammatory responses. In total, 151 proteins were included in the panel and targeted by 243 antibodies (Table S3). This final selection of proteins was based on antibody availability within the HPA project (https://www.proteinatlas.org).[9] One antibody per protein target was included for 78 proteins whereas 73 proteins were targeted by more than one antibody (ranging between two and five antibodies per protein target). This is a common strategy when as in this case using antibodies that were generated against different or same regions of the protein (immunogens with partial to full protein sequence coverage), for example, seen in [10–12]. Quality assessment of these multiple antibodies is defined in the quality control and statistical analysis section below. Most of the antibodies were affinity purified polyclonal rabbit antibodies from the HPA project,[9] with the exception of anti-CHI3L1 (polyclonal goat IgG AF2599 and monoclonal rat IgG_1 MAB25991, R&D Systems, Minneapolis, MN, USA) and anti-HPGDS (monoclonal mouse IgG_1 MAB6487, R&D Systems).

Plasma profiling was performed using a multiplexed bead-based immunoassay based on antibody suspension bead arrays. To create the bead array, antibodies were immobilized onto color-coded magnetic beads (MagPlex-C, Luminex Corp., Austin, TX, USA), as previously described.[11,13] Here, antibodies were diluted to 1.6 μg/mL in MES and coupled to beads that had been activated in *N*-hydroxysulfosuccinimide/1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (Pierce, Thermo Fisher Scientific, Waltham, MA, USA) prior to the coupling. All beads were combined to create a multiplex antibody bead array and successful coupling confirmed using an R-phycoerythrin-conjugated anti-rabbit IgG antibody (Jackson ImmunoResearch, West Grove, PA, USA).

For the assay, plasma samples were distributed in 96-well microtiter plates (Thermo Fisher Scientific) according to a randomized plate layout using a liquid handling device (EVO150, Tecan Group, Männedorf, Switzerland). The layout was designed to give an equal distribution of the three lifestyle groups (anthroposophic, partly anthroposophic, or non-anthroposophic) across the multiple plates, and at the same time, keeping all samples from the four time points of each child within the same plate. Samples were randomized within each plate. Each plate also included a buffer blank in duplicates and a pool of all samples in quadruplicates for determination of technical variation. Using a liquid handler (SELMA, CyBio, Jena, Germany), plasma samples (3 μL) were diluted 1:10 in PBS and labeled with biotin (NHS-PEO4-biotin, Thermo Fisher Scientific) using a 10-fold molar excess over total protein content. The labeled samples were further diluted 1:50 in an assay buffer of 0.5% (w/v) PVA, 0.8% (w/v) PVP, and 0.1% (w/v) casein (all Sigma-Aldrich, St Louis, MO, USA) in PBS supplemented with 10% (v/v) rabbit IgG (Bethyl Laboratories Inc., Montgomery, TX, USA). Samples were heat treated in a water bath at 56°C for 30 min, cooled to room temperature, and combined with the bead array in a 384-well microtiter plate (Greiner Bio-One, Kremsmünster, Austria). After an overnight incubation, unbound proteins were removed by washing the beads in 0.05% Tween-20 in PBS using a plate washer (EL406, Biotek, Winooski, VT, USA). Bound proteins were cross-linked to the antibodies using 0.4% paraformaldehyde in PBS for 10 min. After another wash cycle, a streptavidin-conjugated fluorophore (R-phycoerythrin, Invitrogen, Thermo Fisher Scientific) was used to detect captured proteins followed by read-out of the relative levels of proteins in a FlexMap3D instrument (Luminex Corp.), reporting fluorescent intensity in arbitrary units.

2.4 | Quality control and statistical methods

Statistical analysis and visualization were performed in the programming language and environment for statistical computing R[14,15] and SPSS version 25 (IBM Corp., Armonk, NY, USA). The raw fluorescent intensity data were processed with the R package "MDimNormn" and function "normn_MA" to correct for potential batch effects arising from the multiple microtiter plates.[16] To assess the technical variation, the coefficients of variation (CV) were calculated for each antibody across technical replicates (four per 96-well plate, 16 in total). The median CV per 96-well plate was between 6.5% and 9.5% for the four plates, and 8.6% for the whole assay 384-well plate. There were no plate effects, assessed by Kruskal–Wallis tests. We applied a quality control filtering step, where the data for proteins measured by multiple antibodies were assessed by Spearman rank correlation. Antibody pairs with high agreement (defined by Spearman's Rho ≥ 0.7) were included in the downstream analysis whereas pairs with low correlation (Spearman's Rho < 0.7) were excluded. This resulted in 19 proteins supported by two-three antibodies (Table S4). After this filtering step, a total of 97 proteins (117 antibodies) remained. A merged antibody profile was created for each of the 19 proteins by computing the mean intensity across the antibodies targeting the same protein. In the end, the data from 97 proteins were used in the succeeding statistical analysis (see Figure S1 in the Supporting Information for a functional annotation summary of these proteins).

In the univariate analysis, the non-parametric test of repeated measures, Friedman's test, was used to identify proteins with a time (age) difference, that is, which of the 97 proteins that changed significantly over time (at 6, 12, 24, and 60 months of age). First, all 70 children were compared with respect to differences between the time points in all 97 proteins. Second, stratified analyses were performed for the 97 proteins in which differences between time points were tested within

each lifestyle group (the anthroposophic, partly anthroposophic, or non-anthroposophic children) separately.

To account for multiple testing, the Bonferroni method was used to adjust *P*-values with respect to the number of proteins tested (i.e., the two-sided *P*-values from Friedman's test were multiplied with 97). The Bonferroni method was chosen because it is a conservative test (i.e., low risk of type I error—false positives) which was preferred in this setting with many proteins. Bonferroni correction was applied separately to *P*-values generated from tests in the group of all children and to tests within each lifestyle. All *P*-values presented for the proteins are Bonferroni corrected and corrected *P*-values < 0.05 considered statistically significant. To compare the three lifestyle groups with respect to demographic data, Fisher's exact test (percentage) and Kruskal–Wallis test (median) was used without multiple adjustments. A two-tailed *P*-value < 0.05 was considered statistically significant.

In the multivariate analysis, longitudinal profiles, or trends, in the data were identified with the self-organizing tree algorithm (SOTA) method using the R package "clValid."[17] Log-transformed data were centered (subtracting the mean of a protein from the individual values) and scaled (dividing the centered values by their standard deviation) per protein using the "scale" function in R. For each protein, the median per time point was calculated for all children or for children from each lifestyle group separately and subjected to the SOTA algorithm using default settings and generating between two and nine clusters. The final number of clusters was judged by comparing the added resolution by increasing number of clusters. Data were also visualized in a heatmap, generated with the R package "pheatmap." Data were here log-transformed, centered, and scaled. Hierarchical clustering of rows (corresponding to the individual samples) and columns (corresponding to the proteins) were performed with the Euclidean distance as the distance measure.

Functional annotation of proteins was performed using the PANTHER classification system (version 15, released 2020-02-14 and based on the GO Ontology database from 2020-08-10).[18] Analysis of overrepresented GO biological process terms in the 97 measured proteins used *Homo sapiens* with the whole human proteome as background. The overrepresentation test of significantly altered proteins used the 97 measured proteins as background. FDR < 0.05 was considered significant.

3 | RESULTS

3.1 | Study population

The study population ($n = 70$) constituted of 15 children living in families with anthroposophic, 27 with partly anthroposophic, and 28 with non-anthroposophic lifestyle. Distribution of sex, gestational age, and birth weight were similar between the three lifestyle groups, whereas lifestyle-related exposures as motherst' diet during pregnancy, home delivery, duration of breast feeding, and time points of vaccinations differed significantly (Table 1).

3.2 | Longitudinal development of plasma proteins

To study the general development of plasma proteins across the first 5 years of life irrespective of lifestyle group, we examined the relative levels of 97 proteins in all 70 children using a multiplex bead-based immunoassay in 280 samples. The levels of 68 proteins were significantly altered between any of the time points (ages of 6, 12, 24, and 60 months) (Tables S3 and S5). In the functional classification analysis of these proteins, no biological processes were statistically overrepresented, indicating that the altered proteins represented various functions as those seen for the 97 proteins (Figure S1).

To visualize the trends over time in these significantly changing proteins, we performed SOTA clustering of the 68 proteins for further characterization (Figure 1A). This revealed that the proteins could be grouped in four clusters, or profiles, with characteristic trends. Cluster 1 included 11 proteins that showed similar levels or a slight increase between 6 and 12 months, followed by a steep decrease at 24 months and furthermore increasing levels at 60 months of age. Cluster 2, consisting of 13 proteins, showed a decreasing trend across the time points. Cluster 3 consisted of 21 proteins with a trend that at first showed a decrease in levels, followed by increasing levels at 24 months that at 60 months were approaching the levels in the initial time points. Cluster 4 included 23 proteins with increasing levels across the first 5 years. See Figure 1B for examples of proteins of each cluster, and Table S3 for information of which cluster each protein belonged to.

3.3 | Plasma proteins stratified by lifestyle

We then investigated whether the lifestyle would influence the protein levels during early childhood. First, the SOTA clustering was repeated but now instead stratified by each of the three lifestyles (Figure 2). This again revealed four clusters with similar profiles as the SOTA clustering of the non-stratified data in Figure 1, indicating that proteins developed with the same general pattern regardless of lifestyle. See Table S3 for information of the proteins in each cluster. When comparing the proteins in each SOTA cluster across the three lifestyle groups, only few proteins were unique to a certain lifestyle (Figure 3A). The anthroposophic group had the highest number of unique cluster assignments ($n = 9$): A2M and CSF2 (cluster 2); CR6, FETUB, and SFTPA1/SFTPA2 (cluster 3); and C9, ECM1, PSORS1C2, and SERPINA3 (cluster 4). For the partly anthroposophic group, C4A/C4B (cluster 2), CERS6 (cluster 3), and CHI3L1 and SFTPC (cluster 4) were unique. The non-anthroposophic lifestyle only had one unique protein, APOE, assigned to cluster 2.

Second, to identify proteins that were significantly altered over time within each lifestyle, we performed statistical tests stratified by lifestyle (Tables S3 and S5). Here, in total 59 proteins were altered over time within one or more of the three lifestyle groups. All these 59 proteins were also altered in the group of all children, which additionally had nine more proteins that were not significant within any of the lifestyle groups (Figure 3B). In the group of children with anthroposophic lifestyle, the levels of 30 proteins were significantly altered

TABLE 1 Demographic data on participating children

	Anthroposophic ($N = 15$)	Partly anthroposophic ($N = 27$)	Non-anthroposophic ($N = 28$)	P-value[b]
Mother's diet and antibiotics				
Vegetarian diet during pregnancy	3/15 (20%)	4/26 (15%)	0/25 (0%)	0.060
Organic diet during pregnancy	11/15 (73%)	12/26 (46%)	1/28 (4%)	**1.65×10^{-6}**
Organic diet during breastfeeding	13/15 (87%)	12/26 (46%)	1/28 (4%)	**5.20×10^{-8}**
Antibiotics during pregnancy	2/15 (13%)	6/26 (23%)	5/28 (18%)	0.790
Antibiotics during delivery	1/15 (7%)	0/26 (0%)	1/27 (4%)	0.692
Delivery				
Home delivery	8/15 (53%)	8/27 (27%)	0/28 (0%)	**3.78×10^{-5}**
Caesarean	2/15 (13%)	2/27 (7%)	3/28 (11%)	0.880
Child				
Sex (female)	7/15 (47%)	17/27 (63%)	19/28 (68%)	0.397
Gestational age at birth (w)	40 (38–41)	40 (38–43)	40 (37–41)	0.316
Birth weight (g)	3600 (2950–4480)	3535 (2930–4900)	3648 (2980–4590)	0.979
Having older siblings	11/15 (73%)	17/24 (71%)	19/28 (68%)	1.000
Child's exposures				
Milk formula first week of life	0/14 (0%)	1/26 (4%)	7/28 (25%)	**0.029**
Breastfeeding (any) at 2 months	15/15 (100%)	26/26 (100%)	28/28 (100%)	–
Exclusive breast feeding at 2 months	15/15 (100%)	25/26 (96%)	22/28 (79%)	0.058
Breastfeeding (any) at 6 months	13/15 (87%)	25/27 (93%)	22/28 (81%)	0.535
Exclusive breast feeding at 6 months	6/15 (40%)	8/27 (30%)	1/27 (4%)	**0.005**
Breastfeeding (any) at 12 months	5/15 (33%)	11/26 (42%)	3/28 (11%)	**0.023**
Exclusive breast feeding at 12 months	0/15 (0%)	1/26 (4%)	0/28 (0%)	0.594
Antibiotics before 2 months	0/15 (0%)	0/26 (0%)	1/28 (4%)	1.000
Any vaccine[a] before 6 months	1/15 (7%)	7/27 (26%)	27/28 (96%)	**2.07×10^{-11}**
Any vaccine[a] before 12 months	1/15 (7%)	12/27 (44%)	28/28 (100%)	**2.48×10^{-11}**
Any vaccine[a] before 24 months	6/15 (40%)	17/27 (63%)	28/28 (100%)	**5.00×10^{-6}**
Any vaccine[a] before 60 months	13/15 (87%)	22/27 (82%)	28/28 (100%)	**0.042**
MMR[c] vaccine[a] before 60 months	1/15 (7%)	6/27 (22%)	26/28 (93%)	**1.11×10^{-10}**

Categorical variables: *n*/*N* (%). Continuous variables: median (range).

[a]The Swedish Immunization Programme (for those born 2004–07) offered all children vaccine against diphtheria, tetanus, whooping cough, polio (at 3, 5, 12, and 60 months), *Hemophilus influenzae* type b (at 3, 5, and 12 months), measles, mumps, and rubella (at 18 months and 6 years).[35] The median for the shortest period within a lifestyle group between most recent vaccine prior to blood collection was 1.2 months (at 6 months collection), 4.5 months (at 12 months collection), 3.8 months (at 24 months collection), and 15.1 months (at 60 months collection).

[b]*P*-values from Fisher's exact test for categorical variables and Kruskal–Wallis test for continuous variables.

[c]Measles, mumps, and rubella.

between any of time points, whereas 56 proteins changed in the partly anthroposophic group (Tables S3 and S5). In the non-anthroposophic group, 47 proteins had altered levels (Tables S3 and S5). Out of the 59 proteins that changed in total, 28 proteins were altered independent of lifestyle group as shown in the Venn diagram (Figure 3B and Table S6). No protein was uniquely changed within the anthroposophic group, whereas 11 (A2M, APOC3, CCR6, CD163, CPA3, CRISP3, FETUB, PSORS1C2, SFTPC, SGPL1, and TGFB3) and 2 (HSP90B1 and

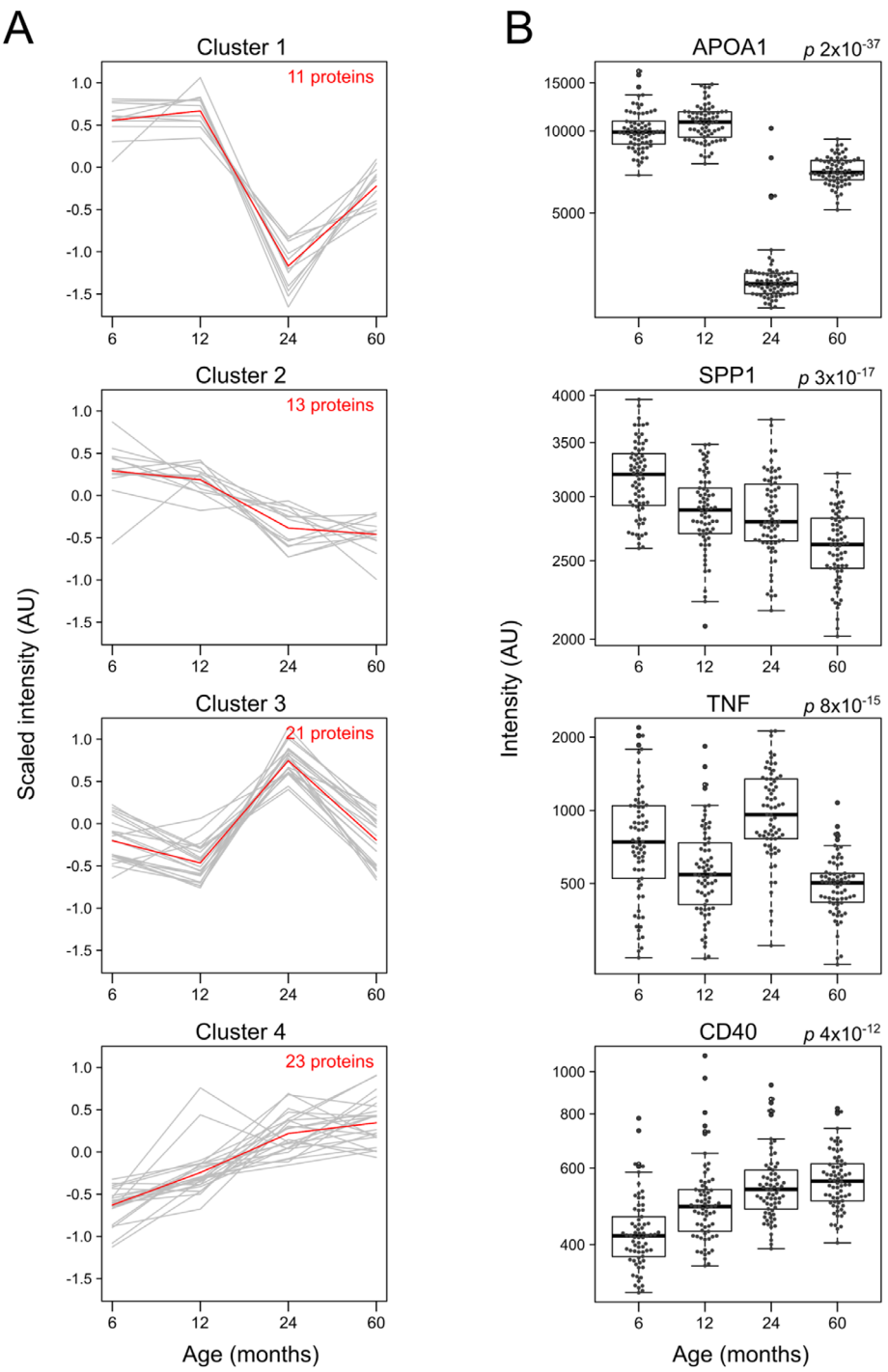

FIGURE 1 Protein trends across time. (A) Clustering was performed using a self-organizing tree algorithm (SOTA) with input from all children ($n = 70$) and the 68 proteins that were significantly altered in Friedman's tests. The trend of individual proteins is shown in gray, mean trend of each cluster in red. (B) Boxplot examples from respective cluster showing the levels of apolipoprotein A1 (APOA1), interleukin 13 (IL13), tumor necrosis factor (TNF), and CD40 molecule (CD40) at the age of 6, 12, 24, and 60 months. Presented *P*-values from Friedman's test were Bonferroni-corrected. AU, arbitrary units

IL13) proteins were unique for the partly anthroposophic and non-anthroposophic groups, respectively (Figure 3B and Table S6).

To further explore the 68 proteins that were altered over time, we performed hierarchical clustering and visualization in a heatmap (Figure 4). This multi-level overview of samples and proteins allowed for identification of potential sub-patterns in the data. On the level of samples, these clustered to some extent based on time point (age) but with an even distribution of the different lifestyles. Each protein was annotated based on the longitudinal profile observed in the SOTA clustering. In the top dendrogram, two main clusters could be observed, one enriched in proteins found in SOTA cluster 1 (blue) and 2 (pink) and the other one enriched in SOTA cluster 3 (yellow) and 4

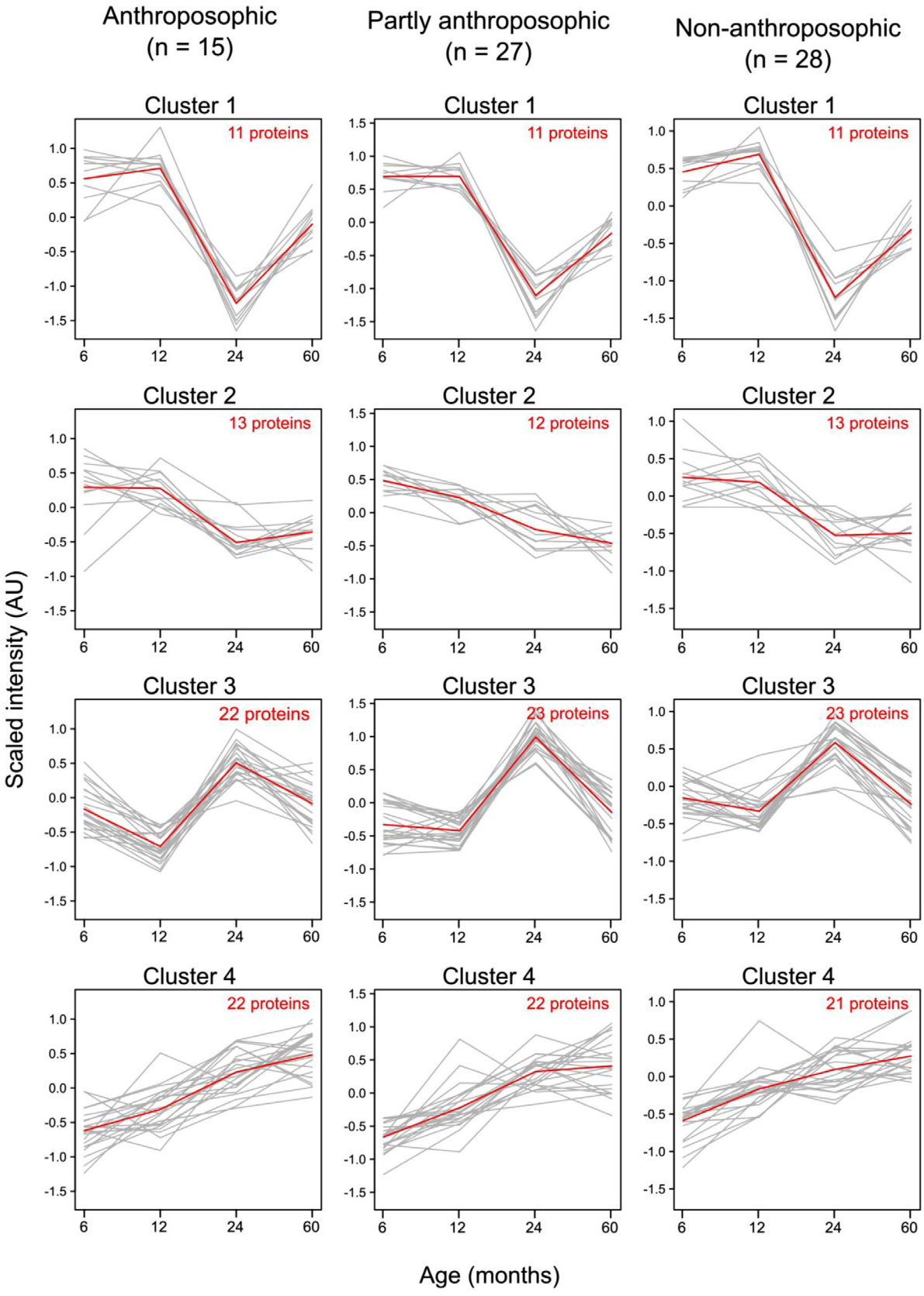

FIGURE 2 Protein trends across time and lifestyle. Clustering using a self-organizing tree algorithm (SOTA) with input from the 68 proteins that were significantly altered in Friedman's tests revealed four clusters. The clustering was performed on children from the respective families separately: anthroposophic, partly anthroposophic, or non-anthroposophic. The trend of individual proteins is shown in gray, mean trend of each cluster in red. AU, arbitrary units

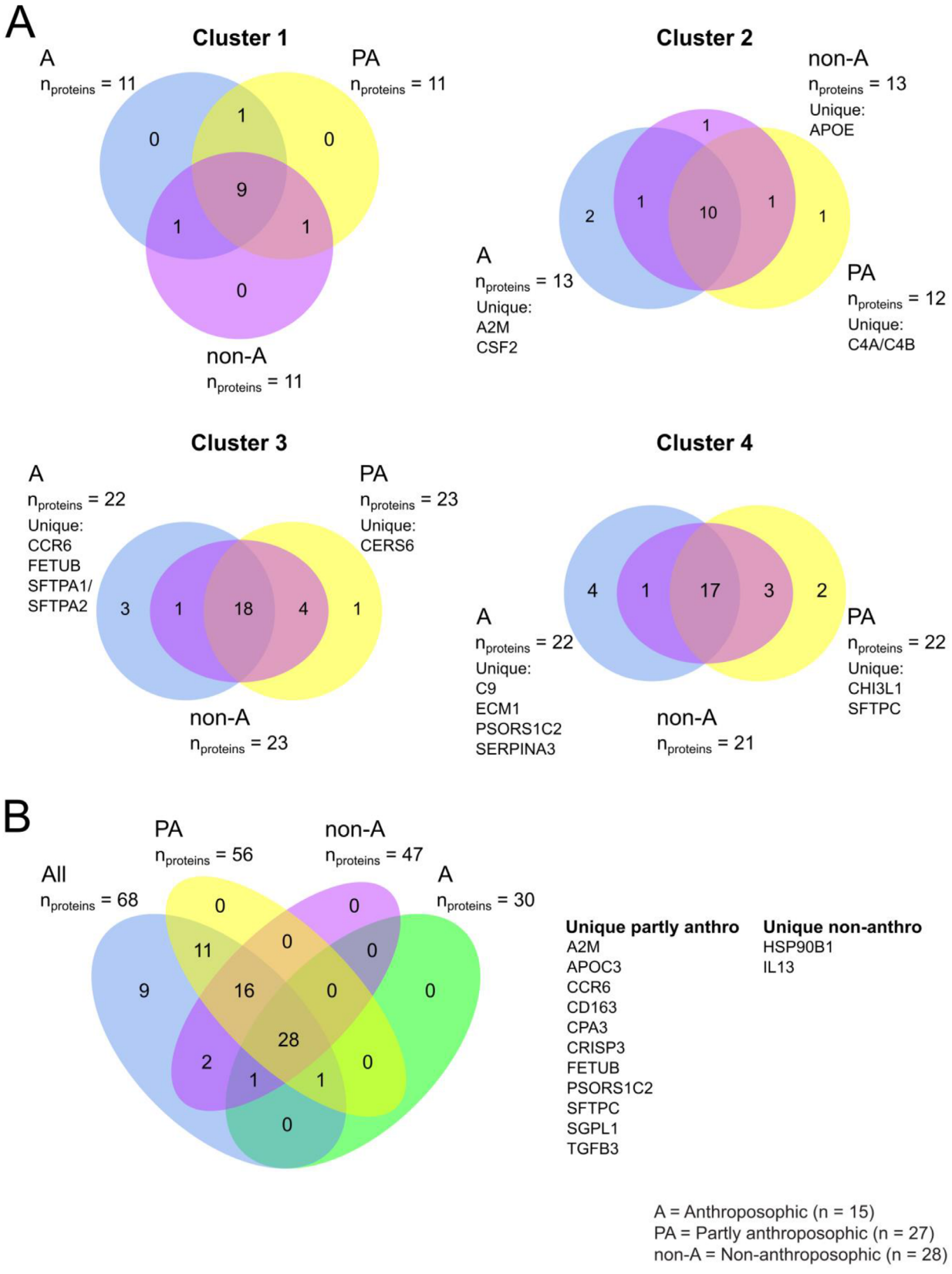

FIGURE 3 Venn diagrams of altered proteins and SOTA clusters. (A) Venn diagrams of the four SOTA clusters in Figure 2, showing the number of proteins that were unique or shared by the lifestyle groups for a certain cluster. Proteins unique for a lifestyle group are shown. (B) Venn diagram showing the number of proteins that were significantly altered over time in all children or altered within children from families with anthroposophic, partly anthroposophic, or non-anthroposophic lifestyle. The names of the proteins being unique for the partly anthroposophic and unique for the non-anthroposophic are shown. No proteins were uniquely altered within the anthroposophic group. See Table S6 for further details

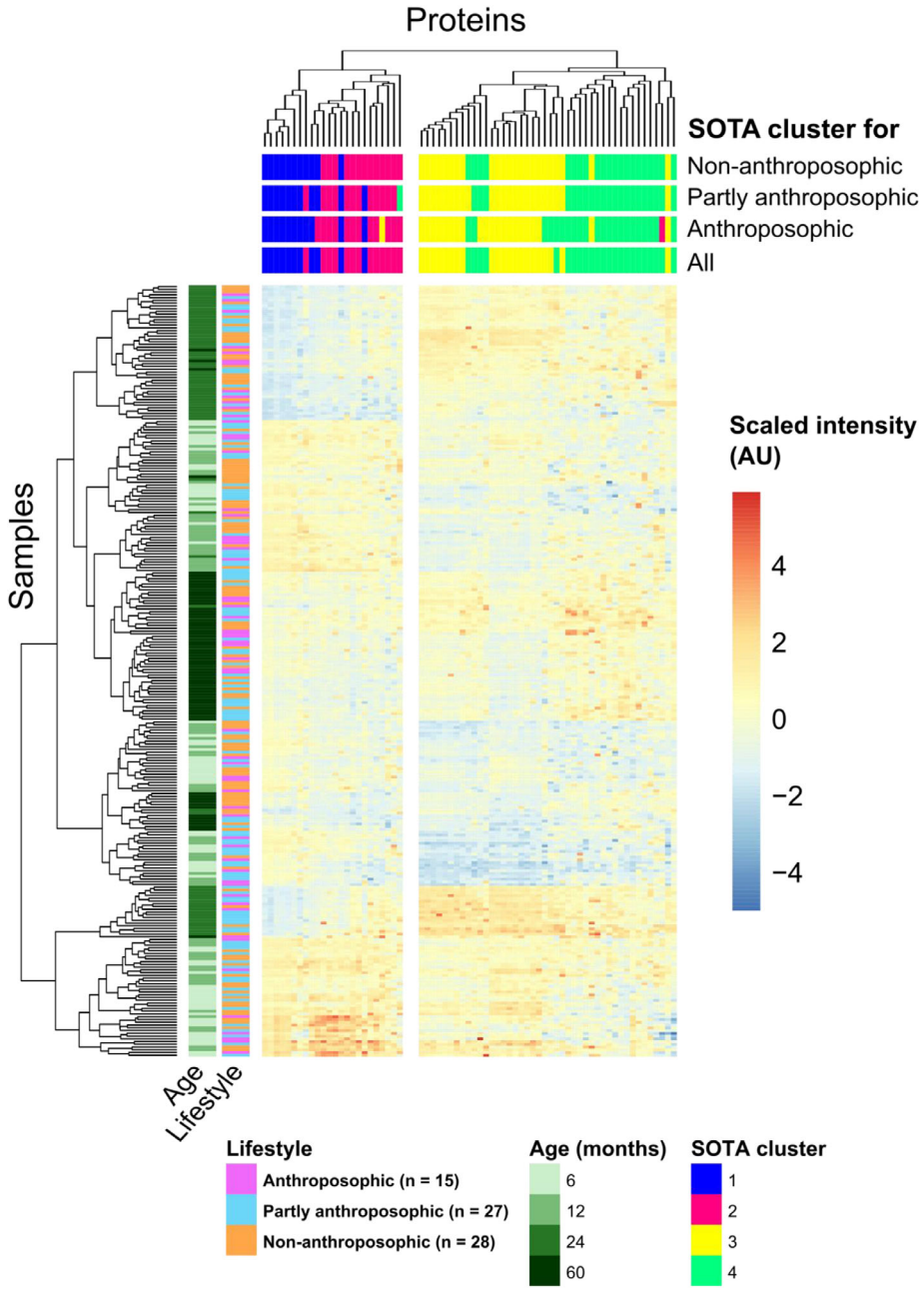

FIGURE 4 Heatmap of all 70 children and the profiles of 68 proteins that were significantly altered using Friedman's tests. Hierarchical clustering on the level of samples (rows) collected at ages 6, 12, 24, and 60 months as well as on the level of proteins (columns). Samples are furthermore annotated based on the lifestyle of the family (anthroposophic = pink, partly anthroposophic = blue, non-anthroposophic = orange) and age (6–60 months corresponding to a gradient of lighter to darker green). Proteins are annotated based on the result from the SOTA clustering performed on data from all children or from each of the lifestyles separately (cluster: 1 = blue, 2 = pink, 3 = yellow, 4 = green). AU, arbitrary units

(green). This indicates that the proteins belonging to SOTA cluster 1 + 2 and SOTA cluster 3 + 4 share similarities in their profiles.

4 | DISCUSSION

This study of a unique longitudinal cohort of children with different lifestyles combined with high-throughput plasma protein profiling showed that a majority (70%) of the 97 plasma proteins analyzed were altered during infancy up to 5 years of age. The temporal changes represented four longitudinal trends of the plasma proteins during development, also following stratification of lifestyle.

There are only few studies of age-associated differences and longitudinal profiles of plasma proteins in pediatric cohorts. Age-related changes of plasma proteins were studied in neonates, infants, children, and adults in two cross-sectional studies.[2,3] In both studies, two age groups (<1 and 1–5 years) match our age interval but as the group of 1–5 years is covered by three individual time points in our data,

a direct comparison is somewhat limited. Nevertheless, in Ignjatovic et al.,[3] age-associated proteins were involved in processes such as the immune response, homeostasis, and hemostasis in agreement with our data. Although only proteins that were lower or higher in adults compared to the younger age groups were reported,[3] we also observed an age-effect for A2M, FGA, SERPINA1, and SERPINA3. Bjelosevic et al.[2] identified proteins involved in hematopoietic development, immune response, and physiological growth to be altered across the age groups. Furthermore, differences between the groups of <1 and 1–5 years were identified for 27 proteins (see Table S4 in [2]). Four of these proteins were included in our analysis and three (C4A, C4B, and C9) were significantly altered over time (Table S3). What becomes evident from comparing our data to cross-sectional studies like these,[2,3] is the added resolution by including multiple time points. This is highlighted by the marked changes that we could observe between the ages of 6, 12, 24, and 60 months but that are averaged and potentially missed when combining subjects of different ages into one group.

We observed four major patterns of longitudinal protein profiles across ages of 6 months to 5 years in the proteins that were significantly different (Figures 1 and 2). In cluster 1 (Figure 1), we found multiple apolipoproteins, proteins that not only play a role in the transport and metabolism of cholesterol and lipids but also in immunity and inflammation, such as the protective role seen for APOA1.[19,20] Cluster 2 showed a decreasing trend over time (Figure 1B) and included secreted phosphoprotein 1 (SPP1), also called osteopontin. SPP1 is involved in several processes like bone formation, wound healing, tumorigenesis, and immunological responses.[21] In agreement with our data, human plasma SPP1 was recently found to have a constant age-related decline from birth until the age of 14 years with the highest levels between 0 and 24 months of age.[22] Cluster 3, which included an increase in protein levels between 12 and 24 months, included proteins that may be induced by the immune stimulation associated with entering preschool, in Sweden usually at 1.5 years of age, as exemplified by TNF.[23] The costimulatory protein CD40, required for activating antigen-presenting cells,[24] showed a cluster 4 profile, which may illustrate a successively maturing immune system. Longitudinal profiling was also performed in Liu et al.,[25] studying the plasma proteome in 10 children between 9 months and 14 years, where the expression of 1747 proteins could be categorized into seven patterns of temporal changes, including a group of no change. Furthermore, they saw an age-effect for 970 of these proteins.[25] More recently, Lietzén et al.[26] found age to be the most dominant factor influencing plasma protein levels during the first 3 years of life in 14 children. Here, the 122 proteins with an age effect represented six temporal patterns. Consistent with our observation, prominent changes in protein levels were seen during the first year of life. We additionally observed continued changes throughout the whole study period up to 5 years of age.

Insights into the longitudinal dynamics of proteins in early preterm life and up to 3 months of life were recently published.[4,27,28] Zhong et al.[28] profiled 448 proteins in serum samples from extremely preterm infants followed from birth up to term-equivalent age. Bennike et al.,[27] studied the plasma proteome over the first week of life in full-term newborns, highlighting developmental changes in components of the complement system including C1QB, C4A, and C9 as also seen in our study, suggesting a continued role of the complement at later ages. Furthermore, in Olin et al.,[4] 267 plasma proteins were profiled at birth and at 1, 4, and 12 weeks of age, with only few children sampled at 6 months of age. Although the sampled timeframe of these studies[4,27,28] is not comparable to ours, they provide observations of protein changes at earlier time points.

Lifestyle and environmental factors are known to influence both health and disease and various exposures may be reflected in changes in the plasma proteome. When we analyzed the effect of lifestyle, 59 of 97 proteins were altered over time within one or more of the lifestyle groups, and nearly half of these (28 out of 59) were altered independent of lifestyle group. There was an influence of lifestyle on a protein level shown by some proteins being changed uniquely within only one of the lifestyle groups (Figure 3B). However, no single protein was identified that only showed changes in any of the three lifestyle groups. This supports the notion that the changes are due to ontogeny and less so due to lifestyle differences (Figure 3B). In addition, the SOTA clustering indicated more lifestyle-related differences (Figure 3A). Interestingly, the highest number of proteins ($n = 9$) with a unique SOTA cluster assignment was found in the anthroposophic group (Figure 3A). However, further studies are required to determine the source of this variability and identify associations with exposures that are specifically related to the families' way of life and to clinical relevance.

A strength in our study is the unique, comparatively large, longitudinal approach in a birth cohort with well characterized different lifestyle groups[5] in which the same children were monitored four times, already from infancy up to 5 years of age. In addition, previously discussed studies,[2,3,25,26] have for method related reasons been relatively limited in sample size. In the present study, we applied an affinity proteomics approach[29] using a bead-based immunoassay with one of its major strengths being that it offers high sample throughput analysis, enabling in this case the large-scale profiling of 280 samples from 70 children. As the method only consumes a few microliters of sample, it is especially useful in settings such as those involving infants and young children, where limited sample volumes may be collected. This method has previously been applied in both pediatric and adult cohorts in other contexts such as childhood asthma and malaria,[12,30] liver and kidney disease,[11,31] nervous system/neurodegenerative disease,[32,33] and cancer.[34]

Although our study included a large number of children, the number of children per lifestyle group was too small to allow for analysis stratified by exposure, such as vaccination and breast feeding or outcome such as sensitization to allergens. An association with sex in the development of plasma proteins cannot be ruled out; however, the proportions of females and males were reasonably balanced, also between the lifestyle groups (Table 1). The method that was applied for plasma profiling used a single binder approach that generated relative levels of proteins,[13] and although this does allow for identification of longitudinal patterns, follow-up with methods for absolute quantification would be of interest to get further insights into the magnitudes of temporal changes and how these differ between proteins as well as for the establishment of reference intervals. Furthermore, multiple

antibodies were used to target 73 of the 151 proteins that were measured. Based on a quality assessment of these antibodies, only supportive antibody pairs were included in the statistical analysis whereas those with low correlation were excluded. Possible explanations for differing results include that the epitopes of the multiple antibodies may differ and/or be affected by post translational modifications and thereby be differentially accessible for antibody recognition and binding as well as off-target binding. Among the 68 proteins that were significantly altered over time, 15 proteins were measured with more than one antibody, adding further technical support and confidence in the measurement of these proteins.

In conclusion, a majority (70%) of the 97 plasma proteins analyzed showed altered plasma levels during infancy up to 5 years of age. This could be visualized as four longitudinal trends, illustrating the plasma proteins being highly variable during infancy and childhood. More than half of the proteins (59 of 97) changed within one or more of the three different lifestyle groups. Among these, almost half (28 of 59) were changed in all lifestyles. Our findings are a contribution in the efforts to understand the early life specific development of the plasma proteome in relation to lifestyle exposures. The importance of different environmental exposures with a global perspective needs to be further investigated including genetic and epigenetic analysis in larger cohorts.

ACKNOWLEDGMENTS

We acknowledge the families participating in the ALADDIN study for their trust and contribution and the ALADDIN team for their involvement in this work, especially nurse and coordinator Margareta Eriksson, medical doctor Fredrik Stenius, and biomedical analysts Monica Nordlund and Carina Wallén. This study was supported by the Swedish Research Council (2012-3011, 2016-1324); the Swedish Research Council for Working Life and Social Research (2006-1630); the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and the Karolinska Institutet and the Karolinska University Hospital (20180307); the Centre for Allergy Research Karolinska Institutet; the ChAMP (Centre for Allergy Research Highlights Asthma Markers of Phenotype) consortium that is funded by the Swedish Foundation for Strategic Research, the Karolinska Institutet, AstraZeneca & Science for Life Laboratory Joint Research Collaboration; and the Vårdal Foundation; the Mjölkdroppen Society (4-1272/2015); the Swedish Asthma and Allergy Association's Research Foundation (2013-0058); the Cancer and Allergy Fund (2018-0923); and the Ekhaga Foundation (2015-85).

CONFLICT OF INTEREST

A. Scheynius is a member in the Joint Steering Committee for the Human Translational Microbiome Program at SciLifeLab/Karolinska Institutet together with Ferring Pharmaceuticals, Switzerland. The rest of the authors declare that they have no relevant conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Maria Mikus https://orcid.org/0000-0001-6560-6124

REFERENCES

1. Boekelheide, K., Blumberg, B., Chapin, R. E., Cote, I., Graziano, J. H., Janesick, A., Lane, R., Lillycrop, K., Myatt, L., States, J. C., Thayer, K. A., Waalkes, M. P., & Rogers, J. M. (2012). Predicting later-life outcomes of early-life exposures. *Environmental Health Perspectives*, *120*(10), 1353–1361. https://doi.org/10.1289/ehp.1204934
2. Bjelosevic, S., Pascovici, D., Ping, H., Karlaftis, V., Zaw, T., Song, X., Molloy, M. P., Monagle, P., & Ignjatovic, V. (2017). Quantitative age-specific variability of plasma proteins in healthy neonates, children and adults. *Molecular & Cellular Proteomics: MCP*, *16*(5), 924–935. https://doi.org/10.1074/mcp.M116.066720
3. Ignjatovic, V., Lai, C., Summerhayes, R., Mathesius, U., Tawfilis, S., Perugini, M. A., & Monagle, P. (2011). Age-related differences in plasma proteins: How plasma proteins change from neonates to adults. *PLoS One*, *6*(2), e17213. https://doi.org/10.1371/journal.pone.0017213
4. Olin, A., Henckel, E., Chen, Y., Lakshmikanth, T., Pou, C., Mikes, J., Gustafsson, A., Bernhardsson, A. K., Zhang, C., Bohlin, K., & Brodin, P. (2018). Stereotypic immune system development in newborn children. *Cell*, *174*(5), 1277–1292.e14. https://doi.org/10.1016/j.cell.2018.06.045
5. Stenius, F., Swartz, J., Lilja, G., Borres, M., Bottai, M., Pershagen, G., Scheynius, A., & Alm, J. (2011). Lifestyle factors and sensitization in children - The ALADDIN birth cohort. *Allergy*, *66*(10), 1330–1338. https://doi.org/10.1111/j.1398-9995.2011.02662.x
6. Alm, J. S., Swartz, J., Lilja, G., Scheynius, A., & Pershagen, G. (1999). Atopy in children of families with an anthroposophic lifestyle. *Lancet*, *353*(9163), 1485–1488. https://doi.org/10.1016/S0140-6736(98)09344-1
7. Swartz, J., Aronsson, B., Lindblad, F., Järnbert-Pettersson, H., Scheynius, A., Pershagen, G., & Alm, J. (2018). Vaccination and allergic sensitization in early childhood - The ALADDIN birth cohort. *EClinicalMedicine*, *4-5*, 92–98. https://doi.org/10.1016/j.eclinm.2018.10.005
8. Swartz, J., Lindblad, F., Arinell, H., Theorell, T., & Alm, J. (2015). Anthroposophic lifestyle and salivary cortisol are associated with a lower risk of sensitization during childhood. *Pediatric Allergy and Immunology*, *26*(2), 153–160. https://doi.org/10.1111/pai.12342
9. Uhlen, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A.-K., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., … Ponten, F. (2015). Tissue-based map of the human proteome. *Science*, *347*(6220), 1260419. https://doi.org/10.1126/science.1260419
10. Ayoglu, B., Chaouch, A., Lochmüller, H., Politano, L., Bertini, E., Spitali, P., Hiller, M., Niks, E. H., Gualandi, F., Pontén, F., Bushby, K., Aartsma-Rus, A., Schwartz, E., Le Priol, Y., Straub, V., Uhlén, M., Cirak, S., 'T Hoen, P. A. C., Muntoni, F., … Al-Khalili Szigyarto, C. (2014). Affinity proteomics within rare diseases: A BIO-NMD study for blood biomarkers of muscular dystrophies. *EMBO Molecular Medicine*, *6*(7), 918–936. https://doi.org/10.15252/emmm.201303724
11. Mikus, M., Drobin, K., Gry, M., Bachmann, J., Lindberg, J., Yimer, G., Aklillu, E., Makonnen, E., Aderaye, G., Roach, J., Fier, I., Kampf, C., Göpfert, J., Perazzo, H., Poynard, T., Stephens, C., Andrade, R. J., Lucena, M. I., Arber, N., … Schuppe-Koistinen, I. (2017). Elevated levels of circulating CDH5 and FABP1 in association with human drug-induced liver injury. *Liver International*, *37*(1), 132–140. https://doi.org/10.1111/liv.13174
12. Reuterswärd, P., Bergström, S., Orikiiriza, J., Lindquist, E., Bergström, S., Andersson Svahn, H., Ayoglu, B., Uhlén, M., Wahlgren, M., Normark, J., Ribacke, U., & Nilsson, P. (2018). Levels of human proteins in plasma

associated with acute paediatric malaria. *Malaria Journal*, *17*(1), 426. https://doi.org/10.1186/s12936-018-2576-y
13. Schwenk, J. M., Gry, M., Rimini, R., Uhlén, M., & Nilsson, P. (2008). Antibody suspension bead arrays within serum proteomics. *Journal of Proteome Research*, *7*(8), 3168–3179. https://doi.org/10.1021/pr700890b
14. Ihaka, R., & Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of computational and Graphical Statistics*, 5, 299–314.
15. R Core Team. (2019). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/
16. Hong, M. - G., Lee, W., Nilsson, P., Pawitan, Y., & Schwenk, J. M. (2016). Multidimensional normalization to minimize plate effects of suspension bead array data. *Journal of Proteome Research*, *15*(10), 3473–3480. https://doi.org/10.1021/acs.jproteome.5b01131
17. Brock, G., Pihur, V., Datta, S., & Datta, S. (2008). clValid: An R package for cluster validation. *Journal of Statistical Software*, *25*(4), 1–22. https://doi.org/10.18637/jss.v025.i04
18. Mi, H., Muruganujan, A., Ebert, D., Huang, X., & Thomas, P. D. (2019). PANTHER version 14: More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Research*, *47*(D1), D419–D426. https://doi.org/10.1093/nar/gky1038
19. Figueroa, D. M., Gordon, E. M., Yao, X., & Levine, S. J. (2019). Chapter 13 - Apolipoproteins as context-dependent regulators of lung inflammation. In R. A. Johnston & B. T. Suratt (Eds.), *Mechanisms and manifestations of obesity in lung disease* (pp. 301–326). Academic Press.
20. Mangaraj, M., Nanda, R., & Panda, S. (2016). Apolipoprotein A-I: A molecule of diverse function. *Indian Journal of Clinical Biochemistry: IJCB*, *31*(3), 253–259. https://doi.org/10.1007/s12291-015-0513-1
21. Sodek, J., Ganss, B., & Mckee, M. D. (2000). Osteopontin. *Critical Reviews in Oral Biology and Medicine: An Official Publication of the American Association of Oral Biologists*, *11*(3), 279–303. https://doi.org/10.1177/10454411000110030101
22. Nourkami-Tutdibi, N., Graf, N., Beier, R., Zemlin, M., & Tutdibi, E. (2020). Plasma levels of osteopontin from birth to adulthood. *Pediatric Blood & Cancer*, *67*(7), e28272. https://doi.org/10.1002/pbc.28272
23. Mehta, A. K., Gracias, D. T., & Croft, M. (2018). TNF activity and T cells. *Cytokine*, *101*, 14–18. https://doi.org/10.1016/j.cyto.2016.08.003
24. Laman, J. D., Claassen, E., & Noelle, R. J. (2017). Functions of CD40 and its ligand, gp39 (CD40L). *Critical Reviews in Immunology*, *37*(2-6), 371–420. https://doi.org/10.1615/CritRevImmunol.v37.i2-6.100
25. Liu, C. - W., Bramer, L., Webb-Robertson, B. - J., Waugh, K., Rewers, M. J., & Zhang, Q. (2017). Temporal profiles of plasma proteome during childhood development. *Journal of Proteomics*, *152*, 321–328. https://doi.org/10.1016/j.jprot.2016.11.016
26. Lietzén, N., Cheng, L., Moulder, R., Siljander, H., Laajala, E., Härkönen, T., Peet, A., Vehtari, A., Tillmann, V., Knip, M., Lähdesmäki, H., & Lahesmaa, R. (2018). Characterization and non-parametric modeling of the developing serum proteome during infancy and early childhood. *Scientific Reports*, *8*(1), 5883. https://doi.org/10.1038/s41598-018-24019-5
27. Bennike, T. B., Fatou, B., Angelidou, A., Diray-Arce, J., Falsafi, R., Ford, R., Gill, E. E., Van Haren, S. D., Idoko, O. T., Lee, A. H., Ben-Othman, R., Pomat, W. S., Shannon, C. P., Smolen, K. K., Tebbutt, S. J., Ozonoff, A., Richmond, P. C., Biggelaar, A. H. J. V. D., Hancock, R. E. W., … Steen, H. (2020). Preparing for life: Plasma proteome changes and immune system development during the first week of human life. *Frontiers in Immunology*, *11*, 2650. https://doi.org/10.3389/fimmu.2020.578505
28. Zhong, W., Danielsson, H., Tebani, A., Karlsson, M. J., Elfvin, A., Hellgren, G., Brusselaers, N., Brodin, P., Hellström, A., Fagerberg, L., & Uhlén, M. (2020). Dramatic changes in blood protein levels during the first week of life in extremely preterm infants. *Pediatric Research*, *89*(3), 604–612. https://doi.org/10.1038/s41390-020-0912-8
29. Stoevesandt, O., & Taussig, M. J. (2012). Affinity proteomics: The role of specific binding reagents in human proteome analysis. *Expert Review of Proteomics*, *9*(4), 401–414. https://doi.org/10.1586/epr.12.34
30. Hamsten, C., Häggmark, A., Grundström, J., Mikus, M., Lindskog, C., Konradsen, J. R., Eklund, A., Pershagen, G., Wickman, M., Grunewald, J., Melén, E., Hedlin, G., Nilsson, P., & Van Hage, M. (2016). Protein profiles of CCL5, HPGDS, and NPSR1 in plasma reveal association with childhood asthma. *Allergy*, *71*(9), 1357–1361. https://doi.org/10.1111/all.12927
31. Neiman, M., Hedberg, J. J., Dönnes, P. R., Schuppe-Koistinen, I., Hanschke, S., Schindler, R., Uhlén, M., Schwenk, J. M., & Nilsson, P. (2011). Plasma profiling reveals human fibulin-1 as candidate marker for renal impairment. *Journal of Proteome Research*, *10*(11), 4925–4934. https://doi.org/10.1021/pr200286c
32. Häggmark, A., Mikus, M., Mohsenchian, A., Hong, M. - G., Forsström, B., Gajewska, B., Barańczyk-Kuźma, A., Uhlén, M., Schwenk, J. M., Kuźma-Kozakiewicz, M., & Nilsson, P. (2014). Plasma profiling reveals three proteins associated to amyotrophic lateral sclerosis. *Annals of Clinical and Translational Neurology*, *1*(8), 544–553. https://doi.org/10.1002/acn3.83
33. Remnestål, J., Just, D., Mitsios, N., Fredolini, C., Mulder, J., Schwenk, J. M., Uhlén, M., Kultima, K., Ingelsson, M., Kilander, L., Lannfelt, L., Svenningsson, P., Nellgård, B., Zetterberg, H., Blennow, K., Nilsson, P., & Häggmark-Månberg, A. (2016). CSF profiling of the human brain enriched proteome reveals associations of neuromodulin and neurogranin to Alzheimer's disease. *Proteomics. Clinical Applications*, *10*(12), 1242–1253. https://doi.org/10.1002/prca.201500150
34. Byström, S., Fredolini, C., Edqvist, P. - H., Nyaiesh, E. - N., Drobin, K., Uhlén, M., Bergqvist, M., Pontén, F., & Schwenk, J. M. (2017). Affinity proteomics exploration of melanoma identifies proteins in serum with associations to T-stage and recurrence. *Translational Oncology*, *10*(3), 385–395. https://doi.org/10.1016/j.tranon.2017.03.002
35. The Public Health Agency of Sweden. (2019). Vaccination programmes. Retrieved from https://www.folkhalsomyndigheten.se/the-public-health-agency-of-sweden/communicable-disease-control/vaccinations/vaccination-programmes/

SUPPORTING INFORMATION

Additional supporting information may be found online https://doi.org/10.1002/prca.202000038 in the Supporting Information section at the end of the article.

How to cite this article: Mikus, M., Järnbert-Pettersson, H., Johansson, C., Nilsson, P., Scheynius, A., & Alm, J. (2021). Protein profiles in plasma: Development from infancy to 5 years of age. *Proteomics Clinical Applications*, e2000038. https://doi.org/10.1002/prca.202000038