

ORIGINAL ARTICLE

WILEY

Pre- and postnatal *Lactobacillus reuteri* treatment alters DNA methylation of infant T helper cells

Anna Forsberg¹ | Johanna Huoman¹ | Simon Söderholm² | Ratnesh Bhai Mehta¹ | Lennart Nilsson³ | Thomas R. Abrahamsson⁴ | Jan Ernerudh⁵ | Mika Gustafsson⁶ | Maria C. Jenmalm¹

¹Division of Inflammation and Infection, Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden

²Division of Molecular Medicine and Virology, Department of Biomedical and Clinical Sciences, Wallenberg Centre for Molecular Medicine, Linköping University, Linköping, Sweden

³Allergy Center, Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden

⁴Crown Princess Victoria's Child and Youth Hospital, Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden

⁵Department of Clinical Immunology and Transfusion Medicine, Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden

⁶Department of Physics, Chemistry and Biology, Bioinformatics, Linköping University, Linköping, Sweden

Correspondence

Johanna Huoman, Division of Inflammation and Infection (II), Department of Biomedical and Clinical Sciences (BKV), House 420, Entrance 68, Level 12, University Hospital Campus, Linköping University, 581 83 Linköping, Sweden.
Email: johanna.huoman@liu.se

Funding information

This study was supported by grants from the Swedish Research Council (grants K2011-56X-21854-01-06, 2015-03807 and 2016-01698), the Heart and Lung Foundation

Abstract

Background: Perinatal childhood exposures, including probiotic supplementation, may affect epigenetic modifications and impact on immune maturation and allergy development. The aim of this study was to assess the effects of pre- and postnatal *Lactobacillus reuteri* supplementation on DNA methylation in relation to immune maturation and allergy development.

Methods: DNA methylation patterns were investigated for allergy-related T helper subsets using a locus-specific method and at a genome-wide scale using the Illumina 450K array. From a randomised, double-blind, placebo-controlled allergy prevention trial with pre- and postnatal probiotic supplementation, CD4+ T helper cells were obtained at birth (from cord blood), and 12 and 24 months of age (total (placebo/probiotics); locus-specific method: CB = 32 (17/15), 12 months = 24 (9/15), 24 months = 35 (15/20); Illumina: CB = 19 (10/9), 12 months = 10 (6/4), 24 months = 19(11/8)).

Results: Comparing probiotics to placebo, the greatest genome-wide differential DNA methylation was observed at birth, where the majority of sites were hypomethylated, indicating transcriptional accessibility in the probiotic group. Bioinformatic analyses, including network analyses, revealed a module containing 91 genes, enriched for immune-related pathways such as chemotaxis, PI3K-Akt, MAPK and TGF- β signalling. A majority of the module genes were associated with atopic manifestations (OR = 1.43, $P = 2.4 \times 10^{-6}$), and a classifier built on this model could predict allergy development (AUC = 0.78, $P = 3.0 \times 10^{-3}$). Pathways such as IFN- γ signalling and T-cell activation were more hypermethylated at birth compared with later in life in both intervention groups over time, in line with DNA methylation patterns in the *IFNG* locus obtained by the locus-specific methodology.

Forsberg and Huoman shared first authorship.

Gustafsson and Jenmalm shared senior authorship.

The peer review history for this article is available at <https://publons.com/publon/10.1111/pai.13240>

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2020 The Authors. *Pediatric Allergy and Immunology* published by John Wiley & Sons Ltd

(grants 20140321 and 20170365), the Ekhaga Foundation, the Åke Wiberg Foundation, the Olle Engkvist Foundation, the Research Council for South-East Sweden, the Swedish Asthma and Allergy Association, the Cancer and Allergy Association, the Foundation Samariten, the Konsul Th C Berghs Foundation for Scientific Research, the University Hospital of Linköping, Sweden, and the Research and Scholarship Foundation in Östergötland, Sweden.

Conclusion: Maternal *L. reuteri* supplementation during pregnancy alters DNA methylation patterns in CD4+ T cells towards enhanced immune activation at birth, which may affect immune maturation and allergy development.

KEYWORDS

allergy, CD4+ T helper cells, DNA methylation, infants, *Lactobacillus reuteri*, postnatal, prenatal, probiotics

1 | INTRODUCTION

Individuals developing allergies already exhibit several differences in their immune system at birth,¹ indicating the importance of foetal life in immune development. Prenatal exposures to maternal dietary and microbial factors may along with genetic susceptibility affect the development of, for example, allergic disease.² Epigenetics constitutes a link between the genetic make-up of the foetus and environmental exposures provided by the mother, ultimately affecting the immune maturation of the new-born *post-partum*.¹⁻⁴ Epigenetic mechanisms modify the accessibility of DNA for transcription and comprise heritable changes in gene expression without altering the DNA sequence. Several processes modify epigenetic memory, for example, by post-translational histone modifications and methylation of DNA CpG dinucleotides.⁵⁻⁷ DNA methylation is generally considered a repressor of transcriptional activity and, in relation to other epigenetic modifications, more stable.^{6,7} In neonates, DNA methylation patterns appear to favour the Th2 skewing of immune responses, which may enable development of allergic disease.^{8,9} During pregnancy, the maternal microbial environment programmes immune development of the child¹ via epigenetic mechanisms, by regulating appropriate maturation of innate immunity as well as T helper (Th) and regulatory T (Treg) cell responses.^{1,2,6} Furthermore, the differentiation of CD4+ T cells such as Th1, Th2, Th17 and Treg cells is epigenetically regulated, also by DNA methylation.⁶

Several methods are available for the analysis of DNA methylation patterns both at a genome-wide level, and on a targeted scale. Utilising a T helper subset locus-specific method, DNA methylation patterns of single CpGs in the gene promoters of four distinct T helper cell subsets, that is Th1, Th2, Th17 and Treg cells, may be investigated. The DNA methylation patterns of these sites accurately predict commitment of naïve T helper cells into the different subsets, and as these sites are not incorporated into epigenome-wide approaches, they remain interesting to study in allergy development.¹⁰ Recently, investigations of epigenome-wide DNA methylation signatures related to allergy development in childhood have revealed associations with sensitisation, as well as food allergies and the development of asthma.^{4,11-13} However, comprehensive studies of the epigenome in CD4+ T cells early in life in allergy-prone children supplemented with probiotics are not available as yet.

In our placebo-controlled, randomised, multicentre allergy prevention trial, participants were supplemented with the probiotic

Key Message

Perinatal probiotic supplementation may affect epigenetic modifications, and epigenome-wide studies of CD4+ T cells in allergy-prone children supplemented with probiotics early in life could reveal mechanistic insights into allergy development. In this study, we assessed the effects of pre- and postnatal *Lactobacillus reuteri* supplementation on DNA methylation in relation to immune maturation and allergy development. The epigenome in CD4+ T cells from children receiving perinatal probiotic treatment was altered towards enhanced immune activation and maturation, particularly at the time of birth, suggesting that probiotics may have immune stimulating effects on the developing foetus.

strain *Lactobacillus reuteri*: mothers from gestational week (GW) 36 until delivery and the infants throughout their first year of life.¹⁴ In this cohort, probiotic supplementation resulted in significantly less IgE-associated eczema at 2 years of age¹⁴ and reduced allergen responsiveness during infancy,¹⁵ but no effect on respiratory allergies at 7 years of age.¹⁶ Other childhood probiotic allergy prevention trials have also shown promising results in preventing IgE-associated eczema, but not asthma or other allergic diseases.¹⁷ Combined pre- and postnatal supplementation seems to give the most beneficial allergy-preventive effects, indicating that epigenetic mechanisms can be important in infant immune maturation.¹⁸ Hence, studying epigenetic modifications during perinatal treatment with *L. reuteri* could give mechanistic insights into the potential allergy-preventive effects of probiotics and its possible regulation by DNA methylation.

The aim of this study was to examine epigenetic patterns in cord blood and peripheral blood CD4+ T helper cells from children in our probiotic intervention study, and to investigate whether these patterns differ depending on supplementation, age or allergic outcome. The DNA methylation patterning of allergy-related T helper cells was investigated utilising a T helper subset locus-specific DNA methylation analysis, and whole-epigenome Illumina 450K DNA methylation analyses were performed to gain insights into the global effects of the intervention. We hypothesised that probiotic supplementation

will lead to hypomethylation, and hence in general to increased transcriptional accessibility, of a subset of immune-related genes in CD4+ T helper cells, both at targeted immunoregulatory CpG sites and at the global scale.

2 | METHODS

For a brief overview of the study, see Figure 1 and Table 1. Detailed information on study design and utilised methods is provided in the Online supplementary material section.

3 | RESULTS

3.1 | T helper subset locus-specific DNA methylation analysis reveals age-dependent differences in immunoregulatory loci

DNA methylation patterns of the T helper cell-associated loci *IL13*, *IL17*, *IFNG* and *FOXP3* did not differ with regard to intervention group or the development of allergic disease. However, age-dependent differences were apparent. The *IFNG* and *IL17* promoters were highly methylated at birth, with significantly less methylation at all later time points (Figure 2A,D) in comparison with at birth. In contrast, the *FOXP3* (Figure 2B) and *IL13* (Figure 2C) loci were significantly more methylated at 12 and 24 months than at birth. In general, the *IL17* locus displayed the highest methylation levels (Figure 2D), and DNA methylation was significantly lower in adults compared with children in all examined loci.

3.2 | Whole-genome DNA methylation patterns show most differences between probiotic and placebo-supplemented infants at birth

Clear differences were observed between probiotic and placebo-supplemented infants in initial multidimensional scaling (MDS) analyses at birth, while fewer differences were seen at 12 and 24 months (Figure 3A-C). Differentially methylated probes (DMPs, defined as mean log₂ methylation difference (MMD) > ± 5% and FDR < 0.1) were extracted, revealing a total of 1629 significant DMPs when comparing the intervention groups at birth, while no or few DMPs were found at 12 and 24 months of age (Figure 3D-F, Tables S1-S4). Among the 1629 DMPs, more than twice as many were hypomethylated than hypermethylated (68% vs 32%). Genomic inflation analyses at birth revealed a slight over-significance in our data (inflation factor of 1.2, bias = -0.003, Figure S1). However, because of the weak inflation, any false positives are avoided through the subsequent *P*-value correction using FDR correction. Thus, no additional correction for inflation was deemed necessary. Additionally, permutation and bootstrap analyses (MATLABs *matte*, 100 000

re-shufflings) confirmed the statistical robustness of the data, as all of the 1629 probes showed nominally significant *P*-values according to both tests. No clear differences were found when comparing children developing any type of atopic manifestations with those who did not, separately at different time points (results not shown). Separate within-group comparisons over time, from birth to 12 and 24 months of age, respectively, were performed (Figure S2, Tables S1, S5-S10) and further mapped to genes. Throughout the first year of life, in genes common to both intervention groups, DMPs were mainly hypermethylated (Figure S3A) and were enriched for pathways such as T-cell activation and IFN- γ signalling (Figure S2B). For the placebo-specific DMPs, hypermethylated DMPs were enriched for pathways of viral immune regulation, whereas genes containing hypomethylated DMPs showed enrichment for mainly non-immune pathways (Figure S2B). Similar results were obtained for the comparison with 2 years of age (Figures S3B and S4).

3.3 | Network approach reveals an immune-associated gene module

As the comparison between the probiotic and placebo groups at birth revealed the most DMPs, these were further mapped to genes, resulting in 1015 genes containing one or more DMPs, of which a majority (65.5%) contained only hypomethylated sites (Figure S5). Enrichment analyses on these genes revealed GO terms mainly related to cell signalling, cell migration and TGF- β signalling (Table S11). In order to further elucidate the potential functional role of the methylation changes in these DMP-containing genes, these were analysed from a network perspective. For this purpose, we computed an intervention module by running four established module identification methods, based on DNA methylation data and publicly available data from the STRING database, and selected a consensus of genes that were predicted by at least two methods and interconnected (see Online supplementary material). The resulting module consisted of 91 genes with 306 direct protein-protein interactions (Figure 4A). Of these module genes, more than half (53 genes) originated from the set of DMP-containing genes in the probiotics vs. placebo comparison at birth, while the remaining 38 constituted novel genes from the interaction network (Figure 4A, Table S12). The module showed overall enrichment for immune-related pathways, such as leucocyte migration and PI3 kinase activity (Figure 4B). Examinations of the network structure of the module identified four large clusters, enriched for chemokine signalling, PI3K-Akt and MAPK signalling, and TGF- β signalling as well as transcription and translation (Figure 4A). Inspecting these 91 genes more thoroughly, 26 genes were highly connected (>10 connections in network, Table S12). Of these highly connected genes, five of the genes that showed more than 5% difference in methylation had previously been associated with the development of allergic outcomes (Table S13). Two of these genes were hypermethylated (*PRKCA*

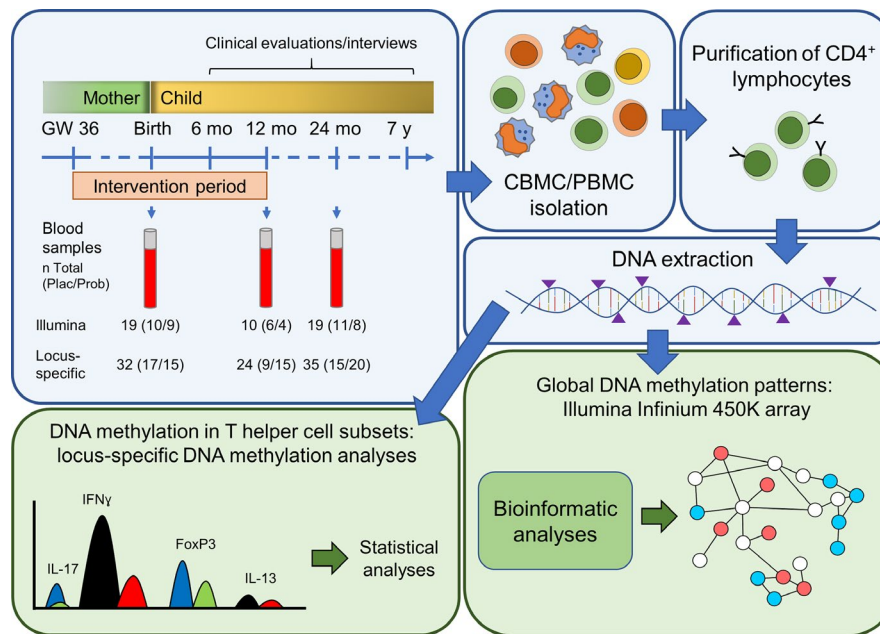


FIGURE 1 Flow scheme for the study. From children in a probiotic intervention study where the mother was treated from GW 36 until birth and the child from birth throughout the first year of life, blood samples were retrieved at birth (cord blood), 12 and 24 mo of age. Thereafter, CBMCs/PBMCs were isolated, and CD4⁺ T cells were purified from these populations. Subsequently, DNA extraction was performed, and utilised for either targeted immune-related (T helper subset locus-specific DNA methylation analysis) or global (Illumina 450K) investigation of DNA methylation patterns, followed by statistical and bioinformatic analyses. The number of samples per method is stated as total (placebo/probiotics) at the different time points. CBMC—cord blood mononuclear cell, GW—gestational week, PBMC—peripheral blood mononuclear cell

and *ITGA4*), and three were hypomethylated (*CCR10*, *OXGR1* and *GNAO1*).

3.4 | Classifier built on module associates with the development of allergic manifestations

In order to further evaluate the module computed on the comparison between probiotics and placebo at birth, the relationship of the module to the development of atopic manifestations was studied. The power of the module (containing 91 genes) to predict which children were going to remain healthy ($n = 17$, $n\text{Prob} = 5$, $n\text{Plac} = 12$, Table S14) or develop atopic manifestations (sensitisation and/or symptoms cumulatively up to 7 years of age, $n = 24$, $n\text{Prob} = 12$, $n\text{Plac} = 12$) was investigated by testing the DMPs of the module ($n = 2734$). These DMPs were extracted from all 91 genes in the network, not only from the genes found to be of interest in the differential DNA methylation analyses. Significantly more genes were differentially methylated between children who later developed atopic manifestations and those who remained healthy ($n = 61$ genes, $\text{OR} = 1.43$, $P = 2.4 \times 10^{-6}$), with most DMPs being hypomethylated (Figure 5A, Table S12). This further supported the possibility of using the module to construct a classifier of the response (see Online supplementary material), which could predict allergy development (as validated by means of leave-one-out cross-validation, average cross-validation $\text{AUC} = 0.78$, $P = 3.0 \times 10^{-3}$ from 100 permutations, Figure 5B).

4 | DISCUSSION

We have demonstrated that perinatal probiotic supplementation affects global DNA methylation patterns of CD4⁺ cells in supplemented infants and that this effect is most pronounced at birth. The identified DMPs were mainly hypomethylated, indicating that the probiotic intervention may mainly modulate the epigenome towards transcriptional accessibility. However, this pattern was not observed at 12 and 24 months of age, highlighting the significance of the prenatal period for immune maturation later in life.^{17,18} As differences in methylation between probiotics and placebo were most pronounced at birth, we proceeded with network analyses to investigate the biological relevance of these findings. The obtained module was enriched for immune-related signalling pathways involving *i.a.* PI3K/MAPK, TGF- β and chemokine signalling. The majority of genes in the module were hypomethylated, suggesting that probiotic intervention may mainly lead to transcriptional activation of these genes, and thereby possibly activate pathways concerning, for example, chemotaxis, T-cell receptor signalling and immunoregulatory functions. The ubiquitous PI3K/MAPK signalling pathways are also involved in signal transduction downstream of the T-cell receptor.¹⁹ In support of our findings, studies on DNA methylation revealed genes such as *MAPK1* and *PI3KD* being differentially methylated in egg-allergic children.¹² The Treg-associated cytokine TGF- β plays a pivotal role in immunoregulatory processes such as induction of allergen

TABLE 1 Descriptive data for the children in the T helper subset locus-specific DNA methylation analyses and Illumina 450K analyses

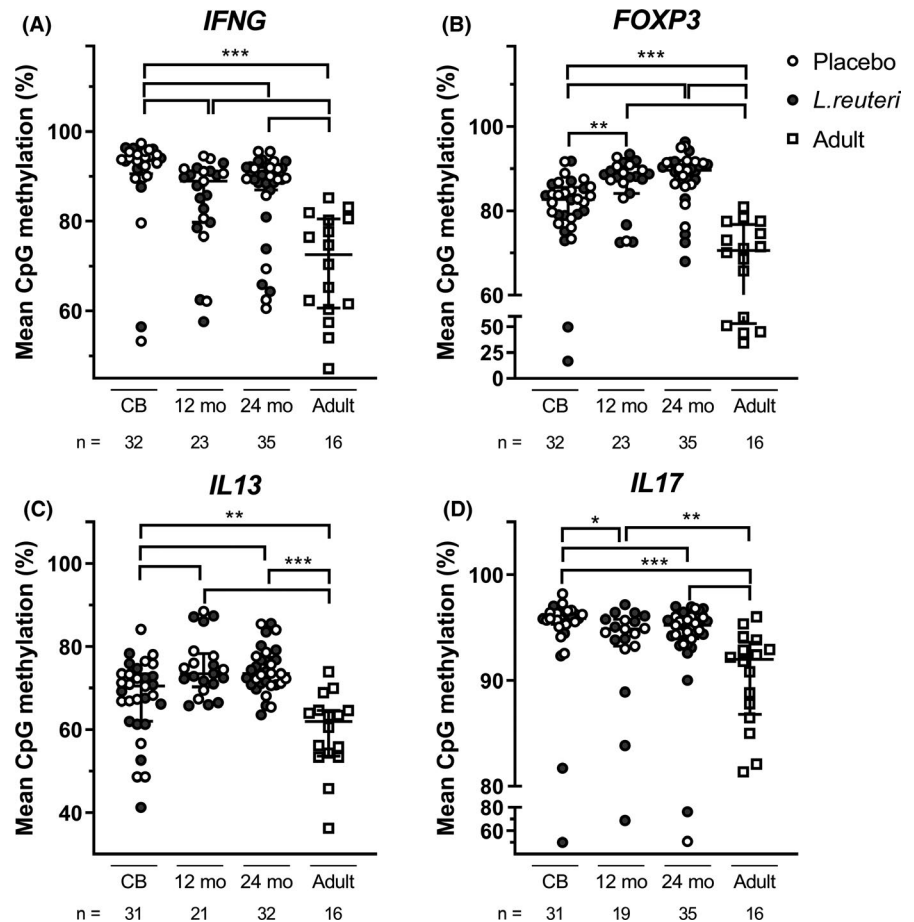
	T helper subset locus-specific DNA methylation analysis (n/N)			Illumina 450K (n/N)		
	CB	12 mo	24 mo	CB	12 mo	24 mo
<i>Lactobacillus reuteri</i>	44 (14/32)	61 (14/23)	57 (20/35)	47 (9/19)	40 (4/10)	42 (8/19)
Boys	31 (10/32) ^a	30 (7/23) ^a	40 (14/35)	42 (8/19)	20 (2/10)	26 (5/19)
First-born	53 (17/32)	52 (12/23)	49 (17/35)	47 (9/19)	50 (5/10)	47 (9/19)
Parental smoking birth	6 (2/32)	4 (1/23)	9 (3/35)	0 (0/19)	0 (0/10)	5 (1/19)
Furred pets	6 (2/32)	22 (5/23)	14 (5/35)	5 (1/19)	40 (4/10) ^a	21 (4/19)
Maternal atopy	78 (25/32)	87(20/23)	91 (32/35) ^a	84 (16/19)	100 (10/10)	90 (17/19)
Breastfeeding 3 m	74 (17/23)	57 (13/23)	69 (24/35)	75 (9/12)	70 (7/10)	79 (15/19)
Breastfeeding 6 m	83 (19/23)	83 (19/23)	91 (32/35)	83 (10/12)	80 (8/10)	95 (18/19)
Antibiotics 0-12 m	17 (4/23)	26 (6/23)	23 (8/35)	17 (2/12)	30 (3/10)	26 (5/19)
Antibiotics 12-24 m	57 (13/23)	52 (12/23)	34 (12/35)	58 (7/12)	50 (5/10)	26 (5/19)
Day care 0-12 m	0 (0/23)	0 (0/23)	3 (1/35)	0 (0/12)	0 (0/10)	0 (0/19)
Day care 12-24 m	91 (21/23)	91 (21/23)	89 (31/35)	92 (11/12)	90 (9/10)	84 (16/19)

Note: ^aSignificantly different from the original study (n = 188), as investigated by chi-square test or Fisher's exact test (for sample sizes with five observations or less).

In comparison with the original study, in the T helper subset locus-specific DNA methylation analyses, significantly fewer boys were included at birth and 12 mo, and significantly more children had atopic mothers at the 24-mo time point. In the Illumina analyses, significantly more children had pets in their homes at the 12-mo time point in comparison with the original study. $P < .05$.

Nine and seven infants were lost to follow-up at 24 mo from birth for T helper subset locus-specific DNA methylation analyses and Illumina analyses, respectively, but were included in the study to investigate effects of maternal probiotic supplementation at birth.

FIGURE 2 Targeted immune-related DNA methylation patterns in T helper cell-associated loci. Age-related changes in specific T helper cell lineage-associated CpG methylation patterns located in proximity to the transcription start site of the respective lineages were analysed by means of T helper subset locus-specific DNA methylation analysis. Mean DNA methylation levels in A *IFNG*, B *FOXP3*, C *IL13* and D *IL17* are displayed at birth (CB), 12 and 24 mo in infants and in a group of adults. Groups were compared using the Kruskal-Wallis test, followed by Dunn's multiple comparison test. FDR-corrected P -values (q -values) are reported as follows, *** $q < 0.001$, ** $q < 0.01$, * $q < 0.05$. CB, cord blood; FDR, false discovery rate



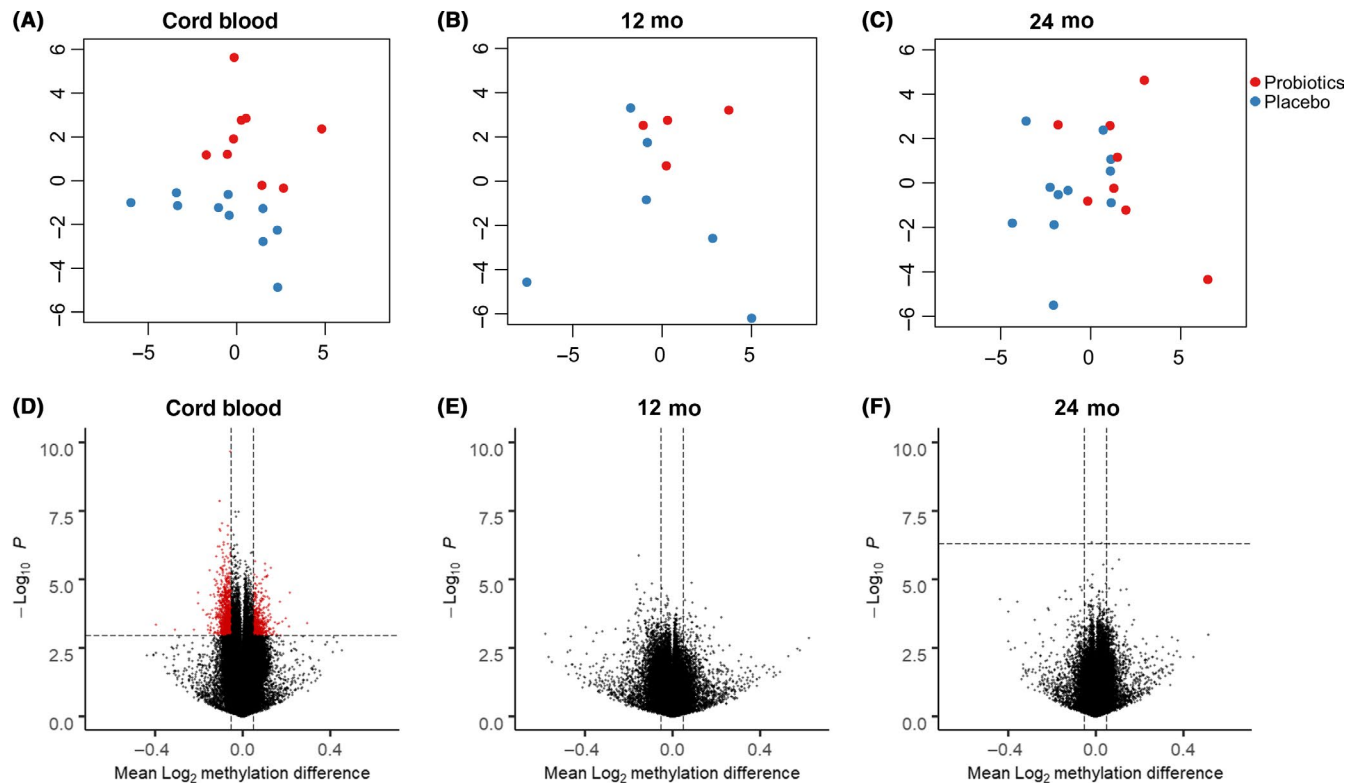


FIGURE 3 Multidimensional scaling (MDS) analysis and volcano plots of differential DNA methylation. MDS was performed on data after filtering, BMIQ normalization and cell-type correction in samples from birth (CB, panel A), 12 months (B) and 24 months (C), respectively. Red dots indicate probiotics, and blue dots denote placebo. Based on the comparison between probiotics and placebo, differential methylation for each probe ($-\log_{10} P$ -values plotted against mean $\log_2$ methylation difference (MMD)) is illustrated as volcano plots in CB at birth (D), and at 12 months (E) and 24 months (F). The differential methylation analysis included fitting a linear model to the obtained data and FDR correction for multiple testing. The horizontal dashed line represents a cut-off of $FDR < 0.1$, and the vertical dashed lines correspond to $MMD < -0.05$ and > 0.05 , respectively. Red dots indicated differentially methylated probes fulfilling both criteria for FDR and MMD described above. BMIQ, beta mixture quantile dilation; CB, cord blood; FDR, false discovery rate; MDS, multidimensional scaling; MMD, mean $\log_2$ methylation difference

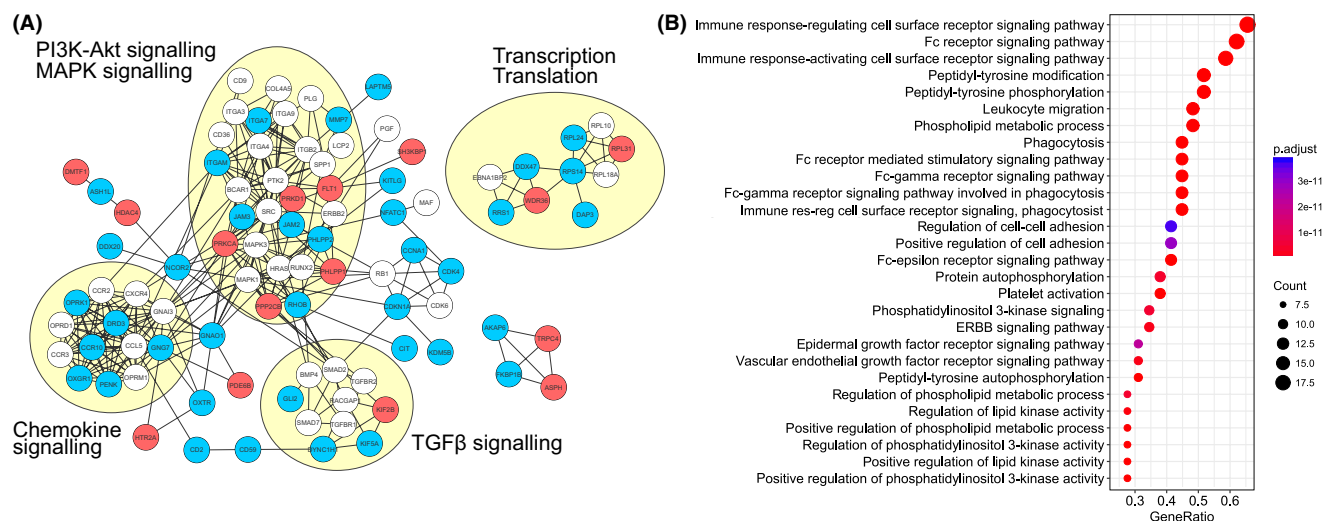
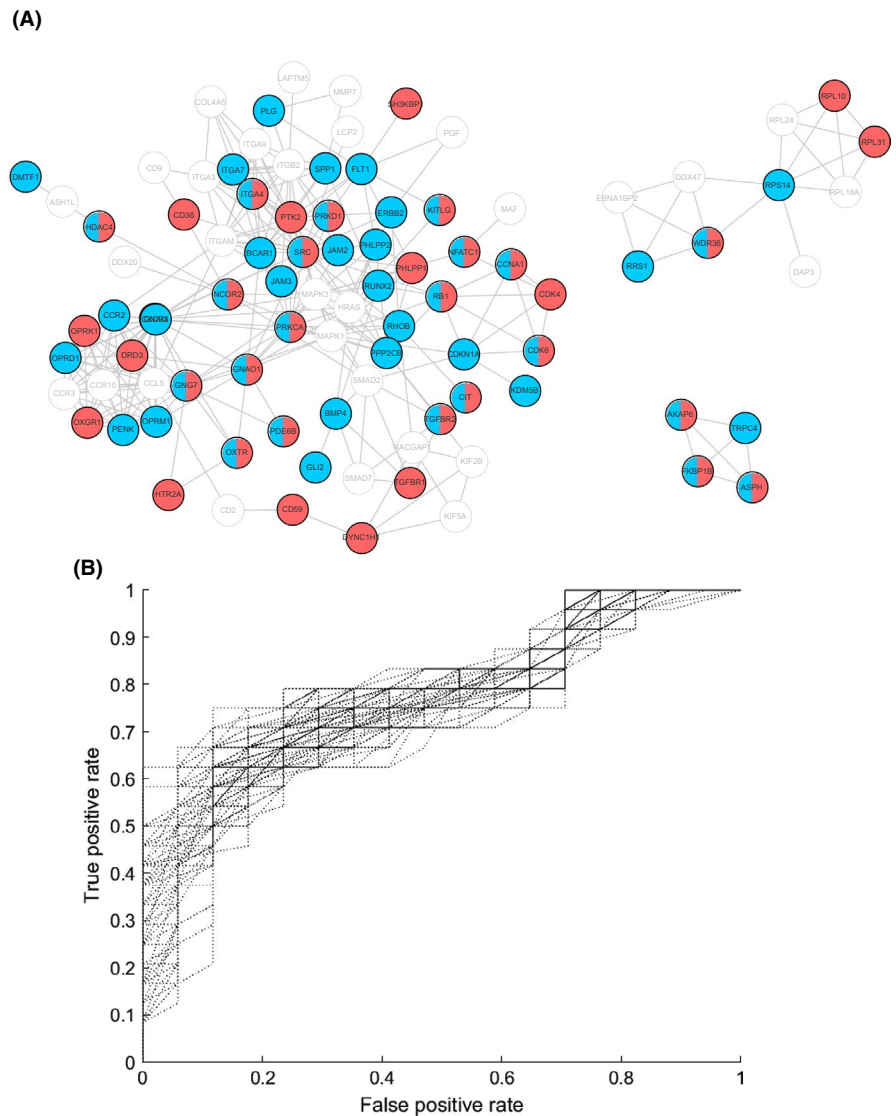


FIGURE 4 Network illustration of module and gene enrichment of module genes. A. Network visualisation of the module identified based on genes containing significant DMPs ($FDR < 0.1$). Nodes represent genes and connecting lines represent protein-protein interactions (STRING combined score > 0.7). Blue nodes correspond to genes containing only hypomethylated DMPs, and red nodes correspond to genes containing only hypermethylated DMPs. Yellow areas display intra-connected subunits, each analysed for biological pathway enrichment using the online tool DAVID. B. Enriched pathways based on all module genes illustrated as a dot plot. The x-axis represents the gene ratio, dot size corresponds to gene count, and dot colour represents adjusted enrichment P -values. DMP, differentially methylated probe; FDR, false discovery rate

FIGURE 5 Visualisation of the module network in relation to the development of atopic manifestations and the allergy predictive power of the module. The left-hand panel A visualises the module network based on relation to the development of atopic manifestations. Module genes that contain probes showing differential methylation when comparing samples from allergic and healthy individuals ($P < .05$) are highlighted in colour. Blue denotes the presence of only hypomethylated probes, red indicates the presence of only hypermethylated probes, and colouring with both blue and red indicates the presence of both hypo- and hypermethylated probes. The right-hand panel B illustrates receiver operating characteristics (ROCs) for the proposed ensemble learner from 100 different repetitions (to account for stochasticity between the runs) when trained only using the probes in the consensus module genes, which were derived from network analysis of probiotics vs. placebo at birth. In each repetition, cross-validated predictions of the learner were calculated to test whether the learner predicted individuals who later developed atopic manifestations or remained healthy up to 7 y of age. The mean area under the curve (AUC) for the 100 learners was 0.78



tolerance, by transducing intracellular signals through SMAD proteins.²⁰ Martino and co-workers found that global DNA methylation patterns in CD4+ T cells from healthy children were enriched in immune-related gene GO terms involved in, for example, TGF- β signalling at 12 months compared with at birth,²¹ suggesting that microbial and environmental exposure throughout the first year of life induced immunoregulatory responses. Intriguingly, network analyses of DNA methylation data in cord blood from children born to asthmatic mothers revealed *SMAD3* as the most connected node in the network,⁴ and Lund et al recently showed that, in children with atopic asthma and a previous history of rhinovirus-induced wheezing, the greatest differential methylation was for the *SMAD3* gene.¹¹ Collectively, these studies corroborate the importance of the TGF- β , PI3K and MAPK pathways for immune maturation throughout childhood. Examinations of the most highly connected genes in the network revealed five differentially methylated ($>\pm 5\%$ MMD) genes that previously had been related to allergy development. The most differentially hypermethylated *PRKCA*, coding for protein kinase C (PKC) alpha, has been

associated with asthma development in genetic studies,^{22,23} and neonatal PKC alpha expression correlates with allergen-stimulated PBMC IL-5 levels at 1 year of age.²⁴ The likewise hypermethylated gene *ITGA4*, coding for integrin subunit alpha 4, is differentially expressed in children developing asthma.²⁵ When coupled with the integrin beta 7 subunit, the heterodimeric protein drives preferential homing of T cells to the gut mucosa.²⁶ The expression of integrin alpha4beta7 in allergen-stimulated CD4+ T cells is also elevated in allergic compared with healthy subjects.²⁷ This is particularly interesting, as the probiotic treatment interacts with the gut mucosa. Furthermore, the CC chemokine receptor 10 (CCR10) was hypomethylated in this study and drives recruitment of T cells to the skin via the epidermally expressed CC chemokine CCL27.²⁸ Significantly increased expression of CCR10 in circulating CD4+ cells was demonstrated in individuals with atopic dermatitis and asthma,²⁹ and in experimental allergic rhinitis, CCR10+ CD4+ memory T cells were present in the nasal mucosa.³⁰ As *L. reuteri* supplementation in this cohort leads to decreased development of IgE-associated eczema,¹⁴ these findings are of notable interest.

The most differentially hypomethylated gene was the G-protein *GNAO1*, with lower expression in PBMCs from allergic rhinitis patients than in healthy subjects.³¹ The likewise hypomethylated GPCR-related gene *OXGR1* binds leukotrienes such as LTE_4 in murine models.³² Other leukotriene receptors, such as *CysLT1*, are upregulated in CD4⁺ T cells of allergic individuals when stimulated with Der p1 allergen³³ and upon treatment with the Th2 cytokine IL-33 in healthy individuals.³⁴ Whether these methylation differences translate to differential protein expression or function remains to be elucidated. It is tempting to speculate that probiotic supplementation leads to improved immune activation at birth compared with non-treated individuals, due to the observed association with epigenome-wide hypomethylation in this study. Furthermore, as food-allergic children have decreased lymphoproliferative responses in infancy, this could possibly be beneficial in relation to allergy development.¹² Recently, a pilot study combining pre- and postnatal *Lactobacillus rhamnosus* GG and *Bifidobacterium lactis* supplementation revealed general hypomethylation of immune-related genes such as *IL-6R*, *IL-5*, *CD38* and *STAT-3*, in treated compared with non-treated children,³⁵ corroborating our findings, suggesting that probiotic supplementation promotes transcriptional accessibility of immune response genes.

To determine whether the module genes were related to the development of atopic manifestations, we employed a linear model of module-related DNA methylation probes with sample time points as co-variables. This approach revealed a number of genes (61 of 91) being differentially methylated in relation to the development of atopic manifestations, encouraging us to build a classifier to evaluate the development of allergic disease. The resulting classifier was able to distinguish healthy individuals from children with atopic manifestations. Although the limited sample size may have affected the accuracy of the classifier, our findings suggest that the created module is not only related to the probiotic intervention, but possibly also related to allergy-preventive effects in our population.

To examine local DNA methylation sites in the promoters of four immunoregulatory T helper cell subsets, that is Th1 (*IFNG*), Th2 (*IL13*), Th17 (*IL17*) and Treg (*FOXP3*) cells, we employed a T helper subset locus-specific DNA methylation analysis method. These four loci are not available in the Illumina 450K panel, and their importance in T helper cell differentiation is established.¹⁰ While neither intervention nor atopic manifestations development had any significant effect on the locus-specific sites, an age-dependent change in methylation status was revealed. In agreement with studies showing lower methylation levels in the Th2-associated *IL13* promoter region in CD4⁺ cells from cord blood,⁸ we observed significantly less methylation of *IL13* at birth, compared with 12 and 24 months of age. Similarly, the *FOXP3* DNA methylation levels increased from birth throughout infancy, whereas the Th1-associated *IFNG* locus methylation decreased from birth onwards, in line with the higher degree of DNA methylation demonstrated in the *IFNG* promoter region in CD4⁺ cells from neonates compared with both children^{9,36} and adults.^{9,36,37} Interestingly, a recent study on cow's milk-allergic children reported increased

DNA methylation in *IL4* and *IL5*, increased demethylation in *FOXP3* and less DNA methylation in *IFNG* and *IL-10* after 12-month treatment with a milk formula containing *Lactobacillus rhamnosus* GG.³⁸ As DNA methylation generally corresponds to decreased transcriptional accessibility,⁶ the decrease in *IFNG* and *FOXP3* methylation may reflect normal immune maturation of Th1 and Treg responses during the first years of life, continuing into adulthood. This is further corroborated by the findings of changes in DNA methylation over time within the treatment groups, revealing a large number of common DMPs, which represent biological processes involved in Th1 immunity and the regulation of viral infections, such as IFN- γ -mediated responses as well as $\text{I}\kappa\text{B}/\text{NF}\kappa\text{B}$ signalling.

There are some limitations to the present study. The limited sample size may have affected our outcomes, and longitudinal analyses were unfortunately not possible. However, we do reveal differences in line with what others have shown^{4,11-12,21} in our cross-sectional analyses, and the means by which we attempted to assure robustness and statistical soundness of the data should at least partly compensate for this. Bootstrap and permutation analyses on the 1629 DMPs at birth revealed significant nominal *P*-values for both methods, indicating that the differential methylation changes are robust. In the downstream network analyses, as a means to further improve the stringency of possible findings, we required that any differences found should be present for a specific probe in at least two of the four module packages tested in order to be considered significant.

To our knowledge, this is the first study to assess the epigenetic developmental trajectory in infant CD4⁺ T cells after pre- and postnatal probiotic supplementation. In conclusion, the epigenome in CD4⁺ T cells from children receiving probiotic treatment pre- and postnatally is altered towards enhanced immune activation and maturation, particularly at the time of birth, suggesting that probiotics have immune-stimulating effects on the developing foetus.

ACKNOWLEDGMENTS

We thank Lena Lindell, Elisabeth Andersson, Linnea Andersson and Eivor Folkesson, Dr Göran Oldaeus and Dr Ted Jacobsson for their brilliant and enthusiastic work guiding the families through the study and all the sampling procedures. We also thank Anne-Marie Fornander for excellent technical assistance and Annelie Castell for ambitious work with the T helper subset locus-specific DNA methylation analyses in the study. Lastly, we would like to thank the SNP&SEQ technology platform at SciLifeLab Uppsala University for their fruitful collaboration.

CONFLICT OF INTEREST

T. Abrahamsson and M. Jenmalm have received funding for clinical trials and honoraria from Biogaia AB for lectures. M. Jenmalm has also received consultant fees and travel support from Nutricia/Danone.

AUTHOR CONTRIBUTIONS

MCJ and TA designed the study. TA and LJN were responsible for the clinical evaluation of the children. JH and AF performed the

experimental work. SS, JH and AF performed statistical analyses and presented the data. MG provided expert knowledge on the performance of network statistical analyses. RBM and JE provided expert knowledge on setting up the experimental method. All authors interpreted and discussed the results. JH, SS and AF drafted the manuscript. All authors contributed to and approved the final draft for publication.

ORCID

Johanna Huoman  <https://orcid.org/0000-0003-2509-2418>

Simon Söderholm  <https://orcid.org/0000-0001-5350-7102>

Lennart Nilsson  <https://orcid.org/0000-0002-5680-6367>

Maria C. Jenmalm  <https://orcid.org/0000-0002-2117-5366>

REFERENCES

- Lockett GA, Huoman J, Holloway JW. Does allergy begin in utero? *Pediatr Allergy Immunol*. 2015;26(5):394-402.
- Sabouchi S, Bollyky J, Nadeau K. Review of environmental impact on the epigenetic regulation of atopic diseases. *Curr Allergy Asthma Rep*. 2015;15(6):33.
- Arshad SH, Karmaus W, Zhang H, Holloway JW. Multigenerational cohorts in patients with asthma and allergy. *J Allergy Clin Immunol*. 2017;139(2):415-421.
- DeVries A, Wlasiuk G, Miller SJ, et al. Epigenome-wide analysis links SMAD3 methylation at birth to asthma in children of asthmatic mothers. *J Allergy Clin Immunol*. 2017;140(2):534-542.
- Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet*. 2003;33(5):245-254.
- Schmidl C, Delacher M, Huehn J, Feuerer M. Epigenetic mechanisms regulating T-cell responses. *J Allergy Clin Immunol*. 2018;142(3):728-743.
- Dor Y, Cedar H. Principles of DNA methylation and their implications for biology and medicine. *Lancet*. 2018;392:777-786.
- Webster RB, Rodriguez Y, Klimecki WT, Vercelli D. The human IL-13 locus in neonatal CD4⁺ T cells is refractory to the acquisition of a repressive chromatin architecture. *J Biol Chem*. 2007;282(1):700-709.
- White GP, Watt PM, Holt BJ, Holt PG. Differential patterns of methylation of the IFN-gamma promoter at CpG and non-CpG sites underlie differences in IFN-gamma gene expression between human neonatal and adult CD45RO⁺ T cells. *J Immunol*. 2002;168(6):2820-2827.
- Janson PC, Linton LB, Bergman EA, et al. Profiling of CD4⁺ T cells with epigenetic immune lineage analysis. *J Immunol*. 2011;186(1):92-102.
- Lund RJ, Osmala M, Malonzo M, et al. Atopic asthma after rhinovirus-induced wheezing is associated with DNA methylation change in the SMAD3 gene promoter. *Allergy*. 2018;73(8):1735-1740.
- Martino D, Neeland M, Dang T, et al. Epigenetic dysregulation of naive CD4⁺ T-cell activation genes in childhood food allergy. *Nat Commun*. 2018;9(1):3308.
- Peng C, Meel ER, Cardenas A, et al. Epigenome-wide association study reveals methylation pathways associated with childhood allergic sensitization. *Epigenetics*. 2019;14:445-466.
- Abrahamsson TR, Jakobsson T, Böttcher MF, et al. Probiotics in prevention of IgE-associated eczema: a double-blind, randomized, placebo-controlled trial. *J Allergy Clin Immunol*. 2007;119(5):1174-1180.
- Forsberg A, Abrahamsson TR, Björkstén B, Jenmalm MC. Pre- and post-natal *Lactobacillus reuteri* supplementation decreases allergen responsiveness in infancy. *Clin Exp Allergy*. 2013;43(4):434-442.
- Abrahamsson TR, Jakobsson T, Björkstén B, Oldaeus G, Jenmalm MC. No effect of probiotics on respiratory allergies: a seven-year follow-up of a randomized controlled trial in infancy. *Pediatr Allergy Immunol*. 2013;24(6):556-561.
- Cuello-Garcia CA, Brozek JL, Fiocchi A, et al. Probiotics for the prevention of allergy: a systematic review and meta-analysis of randomized controlled trials. *J Allergy Clin Immunol*. 2015;136(4):952-961.
- Jenmalm MC. The mother-offspring dyad: microbial transmission, immune interactions and allergy development. *J Intern Med*. 2017;282(6):484-495.
- Gaud G, Lesourne R, Love PE. Regulatory mechanisms in T cell receptor signalling. *Nat Rev Immunol*. 2018;18(8):485-497.
- Palomares O, Martín-Fontecha M, Lauener R, et al. Regulatory T cells and immune regulation of allergic diseases: roles of IL-10 and TGF- β . *Genes Immun*. 2014;15(8):511-520.
- Martino D, Maksimovic J, Joo JH, Prescott SL, Saffery R. Genome-scale profiling reveals a subset of genes regulated by DNA methylation that program somatic T-cell phenotypes in humans. *Genes Immun*. 2012;13(5):388-398.
- Murphy A, Tantisira KG, Soto-Quirós ME, et al. PRKCA: a positional candidate gene for body mass index and asthma. *Am J Hum Genet*. 2009;85(1):87-96.
- Melén E, Himes BE, Brehm JM, et al. Analyses of shared genetic factors between asthma and obesity in children. *J Allergy Clin Immunol*. 2010;126(3):631.
- Prescott SL, Irvine J, Dunstan JA, Hii C, Ferrante A. Protein kinase Czeta: a novel protective neonatal T-cell marker that can be up-regulated by allergy prevention strategies. *J Allergy Clin Immunol*. 2007;120(1):200-206.
- Wang C, Li H, Cao L, Wang G. Identification of differentially expressed genes associated with asthma in children based on the bioanalysis of the regulatory network. *Mol Med Rep*. 2018;18(2):2153-2163.
- Gorfu G, Rivera-Nieves J, Ley K. Role of beta7 integrins in intestinal lymphocyte homing and retention. *Curr Mol Med*. 2009;9(7):836-850.
- Eigenmann PA, Tropia L, Hauser C. The mucosal adhesion receptor alpha4beta7 integrin is selectively increased in lymphocytes stimulated with beta-lactoglobulin in children allergic to cow's milk. *J Allergy Clin Immunol*. 1999;103(5 Pt 1):931-936.
- Xiong N, Fu Y, Hu S, Xia M, Yang J. CCR10 and its ligands in regulation of epithelial immunity and diseases. *Protein Cell*. 2012;3(8):571-580.
- Hijnen D, Nijhuis E, Bruin-Weller M, et al. Differential expression of genes involved in skin homing, proliferation, and apoptosis in CD4⁺ T cells of patients with atopic dermatitis. *J Invest Dermatol*. 2005;125(6):1149-1155.
- Nagakubo D, Yoshie O, Hirata T. Upregulated CCL28 expression in the nasal mucosa in experimental allergic rhinitis: implication for CD4⁺ memory T cell recruitment. *Cell Immunol*. 2016;302:58-62.
- Liu Z, Yelverton RW, Kraft B, Tanner SB, Olsen NJ, Aune TM. Highly conserved gene expression profiles in humans with allergic rhinitis altered by immunotherapy. *Clin Exp Allergy*. 2005;35(12):1581-1590.
- Kanaoka Y, Maekawa A, Austen KF. Identification of GPR99 protein as a potential third cysteinyl leukotriene receptor with a preference for leukotriene E4 ligand. *J Biol Chem*. 2013;288(16):10967-10972.
- Thivierge M, Turcotte S, Rola-Pleszczynski M, Stankova J. Enhanced cysteinyl-leukotriene type 1 receptor expression in T cells from house dust mite-allergic individuals following stimulation with Der p. *J Immunol Res*. 2015;2015:384780.
- Boudaud M, Turcotte S, Stankova J, Rola-Pleszczynski M. IL-33 up-regulates cysteinyl leukotriene receptor type 1 expression in human peripheral blood CD4⁺ T lymphocytes. *J Immunol*. (Baltimore, MD: 1950). 2018;201(9):2787-2798.
- Vähämäki S, Laiho A, Lund R, Isolauri E, Salminen S, Laitinen K. The impact of probiotic supplementation during pregnancy on DNA

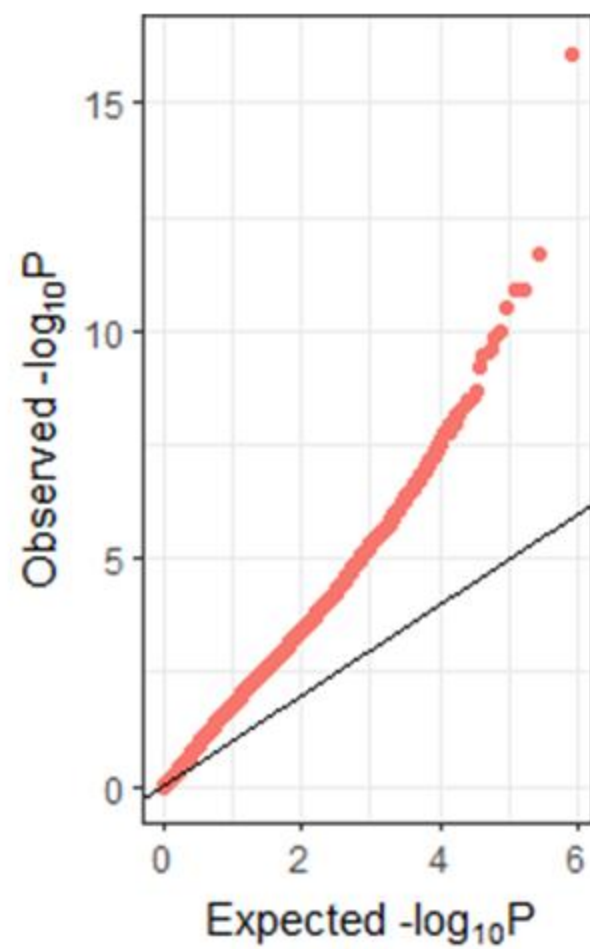
- methylation of obesity-related genes in mothers and their children. *Eur J Nutr.* 2019;58(1):367-377.
36. White GP, Hollams EM, Yerkovich ST, et al. CpG methylation patterns in the IFN γ promoter in naive T cells: variations during Th1 and Th2 differentiation and between atopics and non-atopics. *Pediatr Allergy Immunol.* 2006;17(8):557-564.
37. Jacoby M, Gohrbandt S, Clausse V, Brons NH, Muller CP. Interindividual variability and co-regulation of DNA methylation differ among blood cell populations. *Epigenetics.* 2012;7(12):1421-1434.
38. Paparo L, Nocerino R, Bruno C, et al. Randomized controlled trial on the influence of dietary intervention on epigenetic mechanisms in children with cow's milk allergy: the EPICMA study. *Sci Rep.* 2019;9(1):2828.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Forsberg A, Huoman J, Söderholm S, et al. Pre- and postnatal *Lactobacillus reuteri* treatment alters DNA methylation of infant T helper cells. *Pediatr Allergy Immunol.* 2020;00:1-10. <https://doi.org/10.1111/pai.13240>

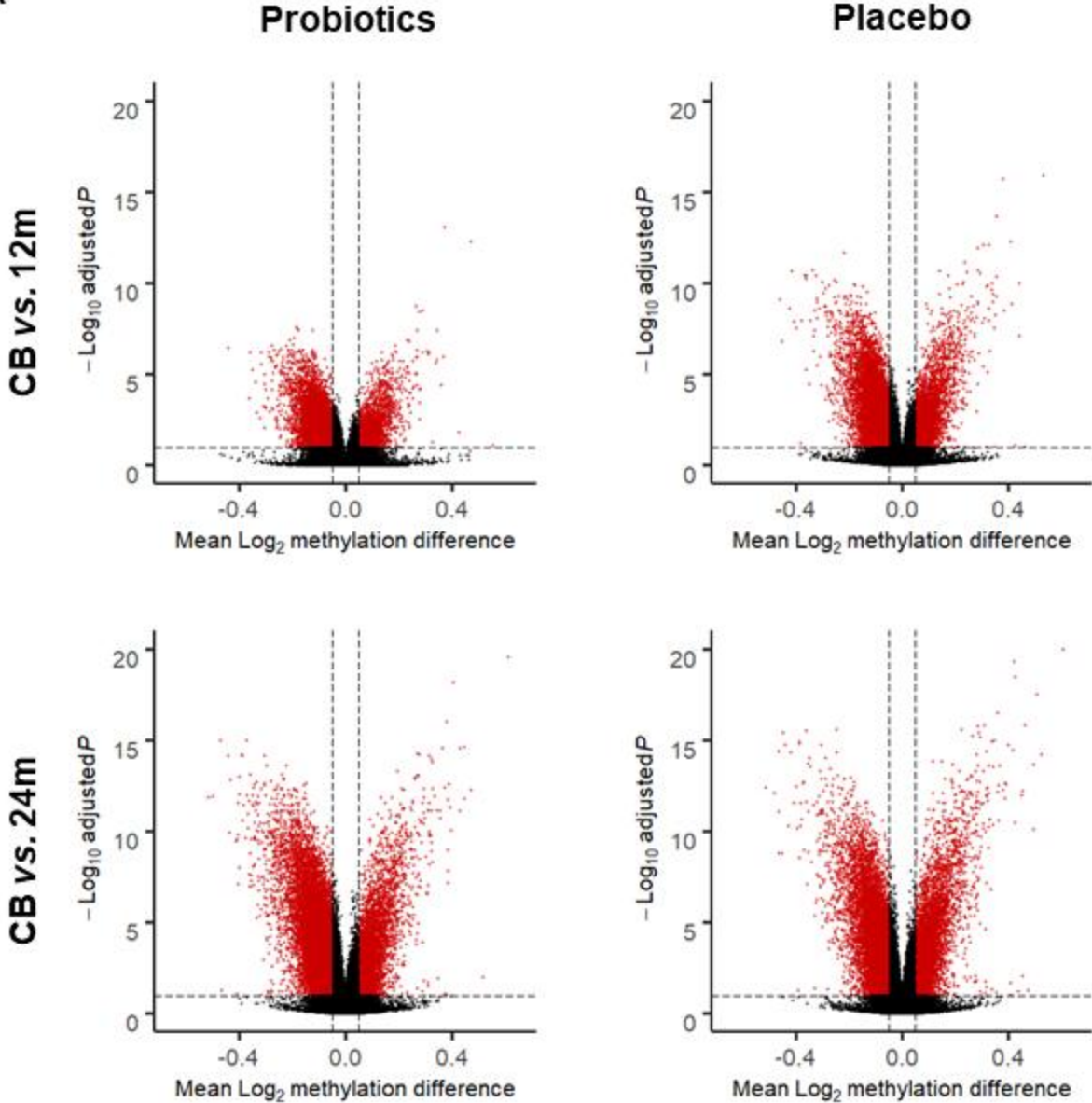
Supplementary Figure 1.



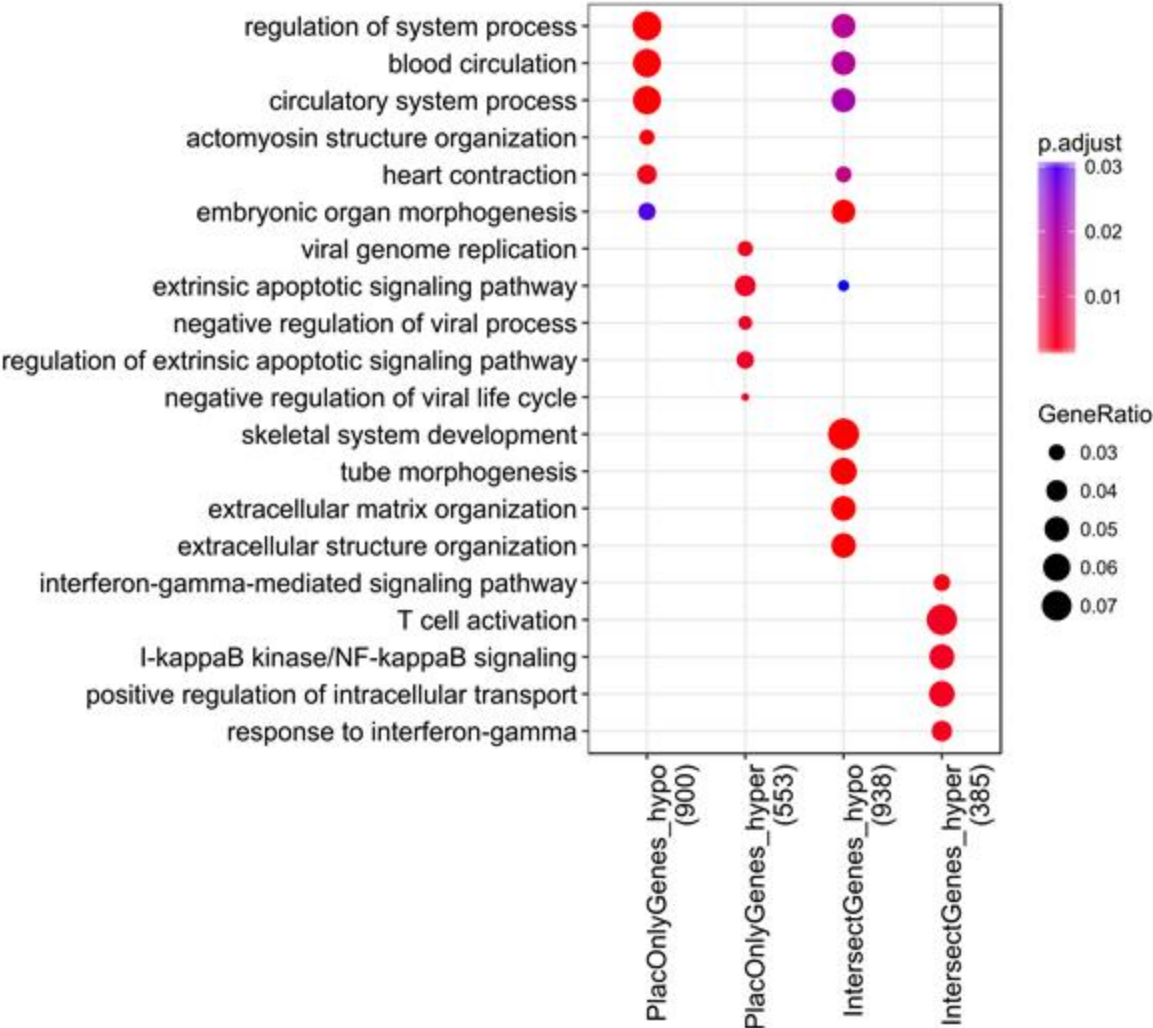
Estimated bias: -0.003
Estimated inflation: 1.2

Supplementary Figure 2.

A

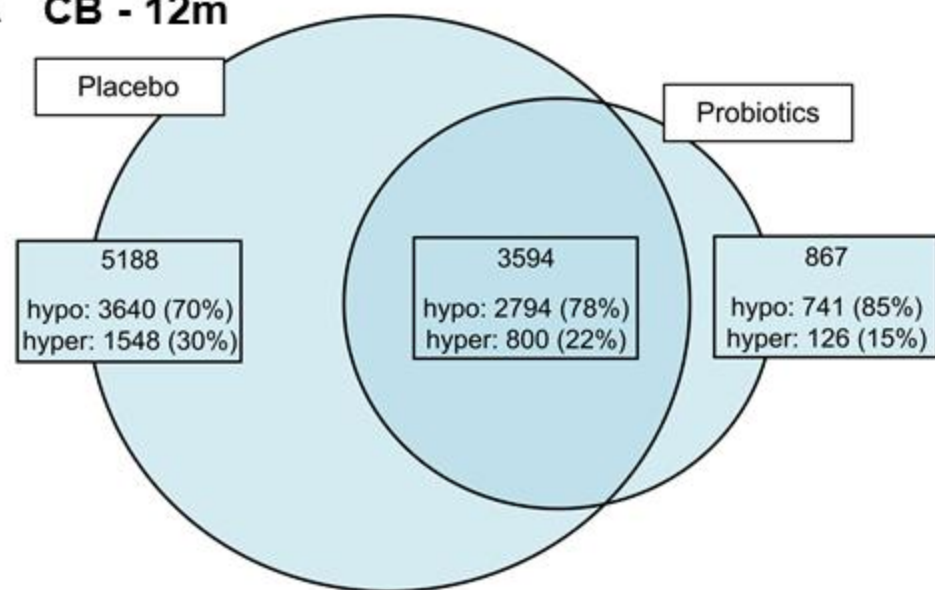


B

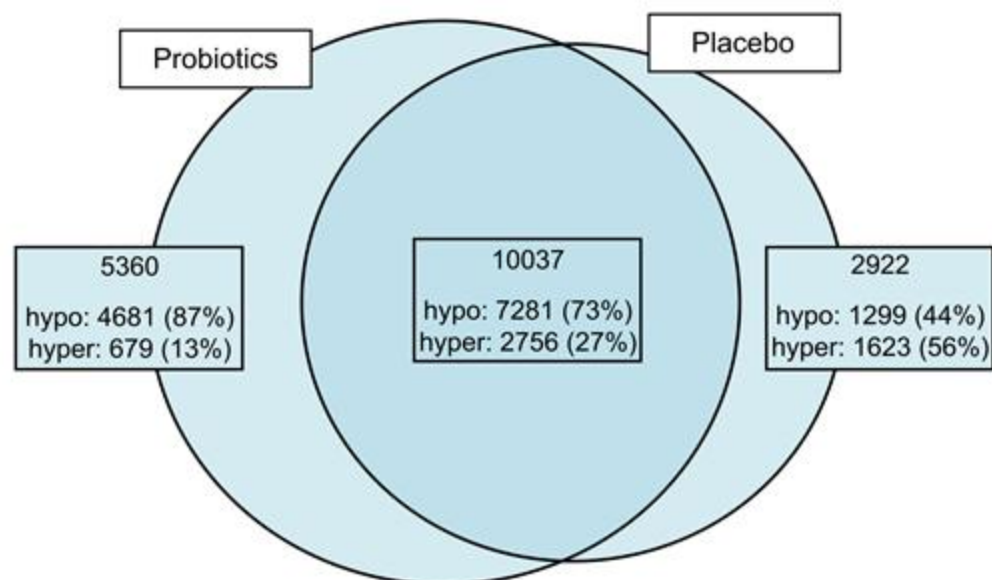


Supplementary Figure 3.

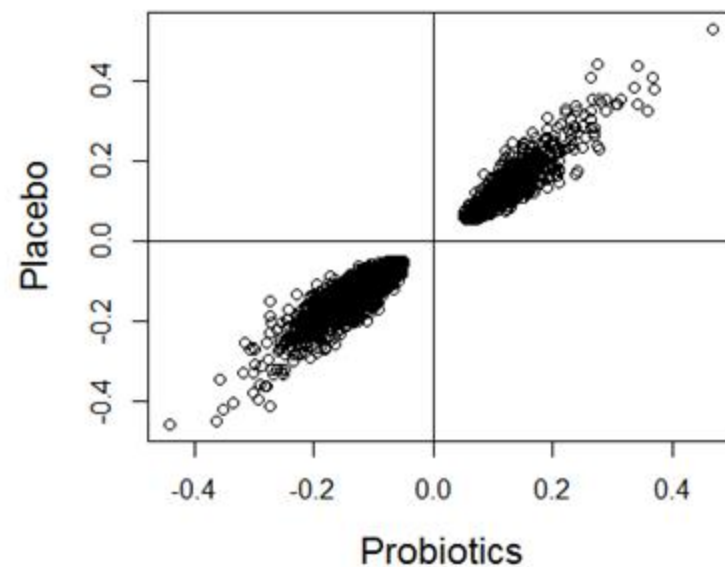
A CB - 12m



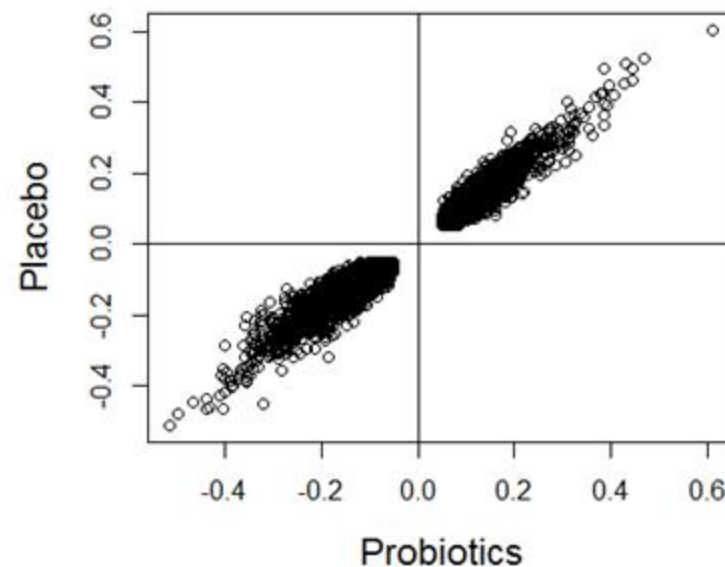
B CB - 24m



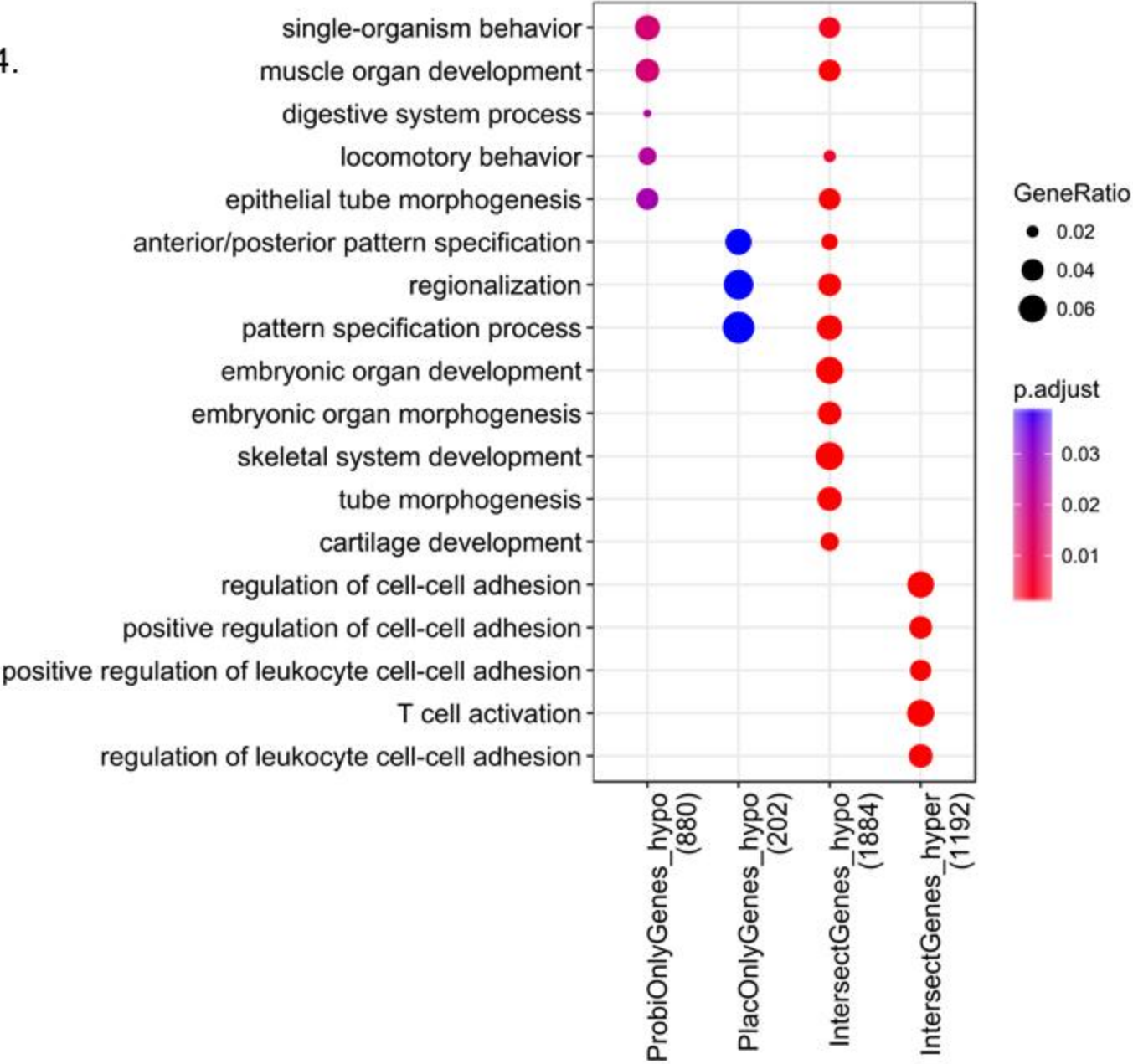
C CB - 12m



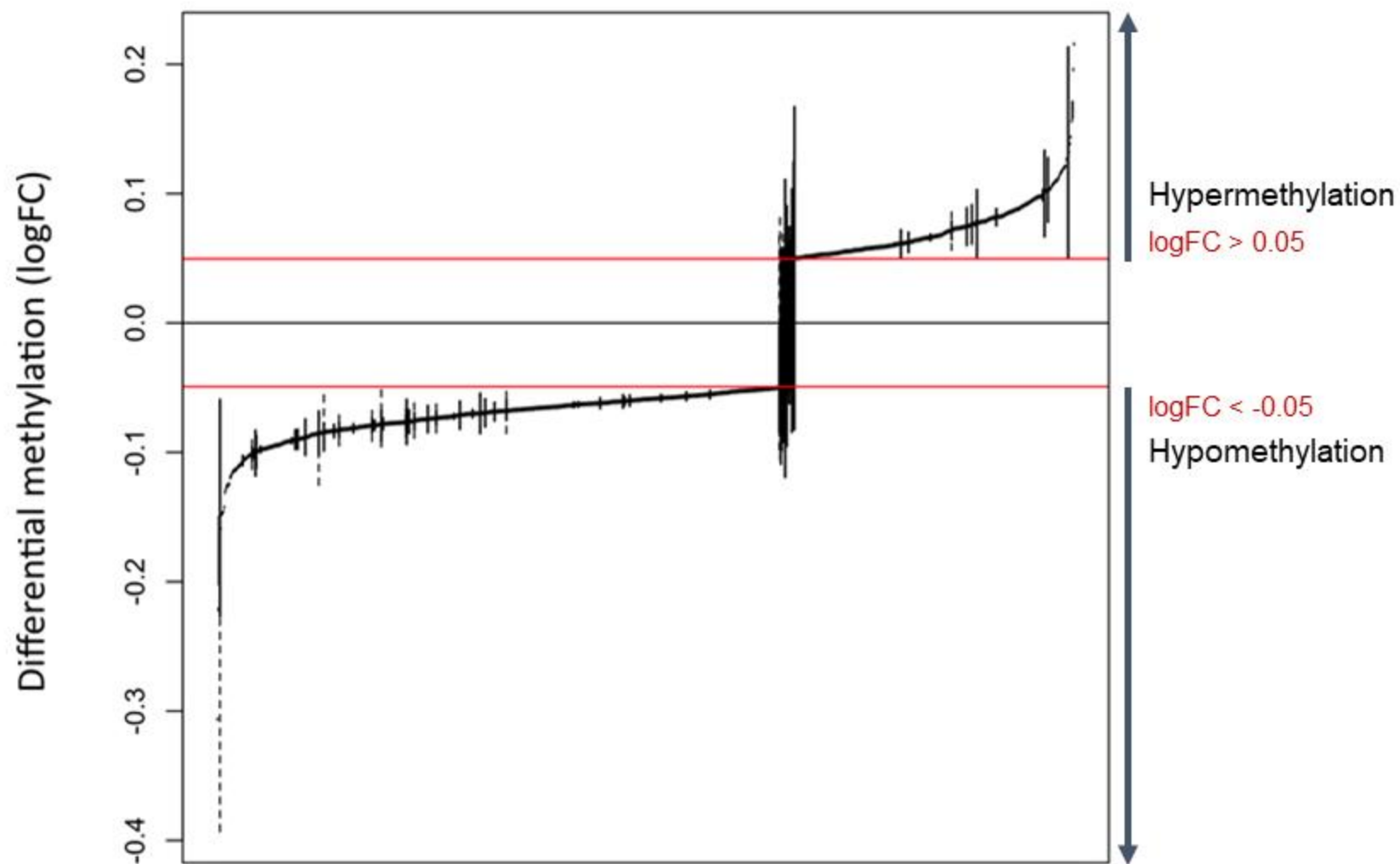
D CB - 24m



Supplementary Figure 4.

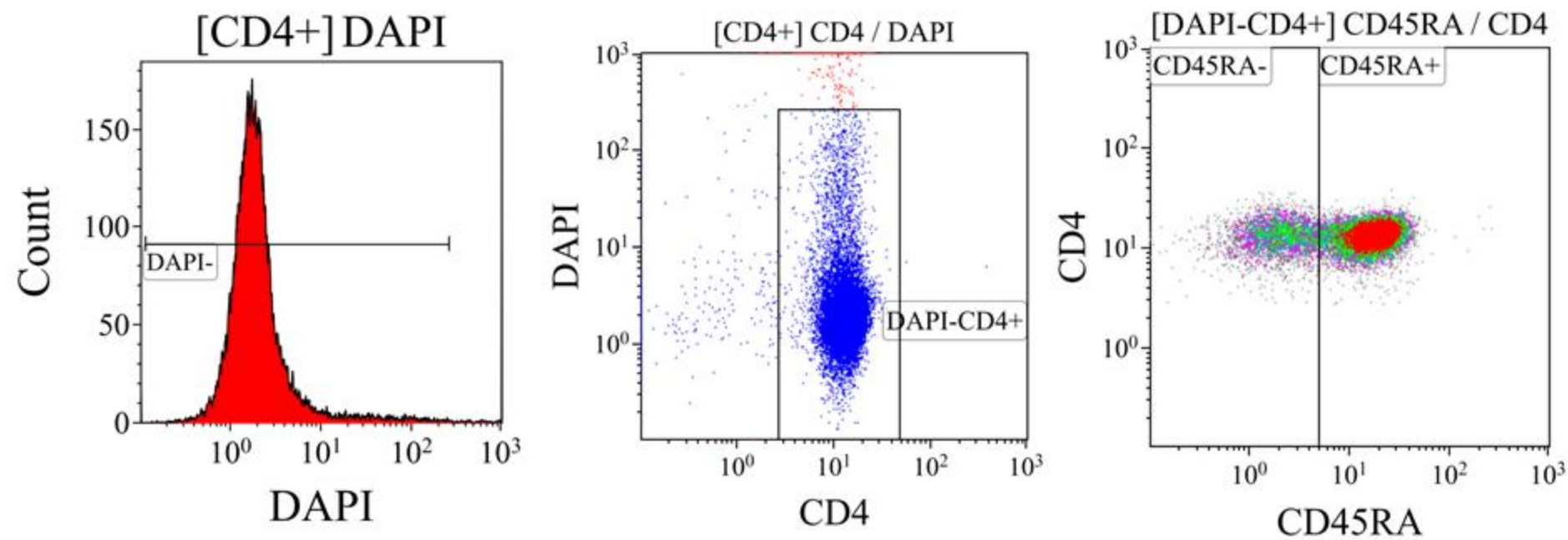


Supplementary Figure 5.



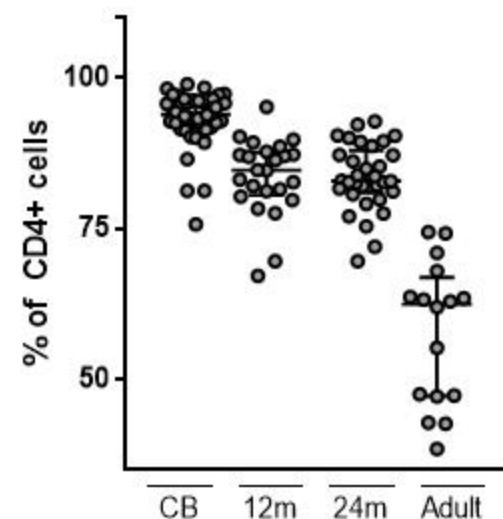
Methylation	# of genes (%)
Hyper	331 (32.6)
Hypo	665 (65.5)
Both	19 (1.9)

Supplementary Figure 6.



Supplementary Figure 7.

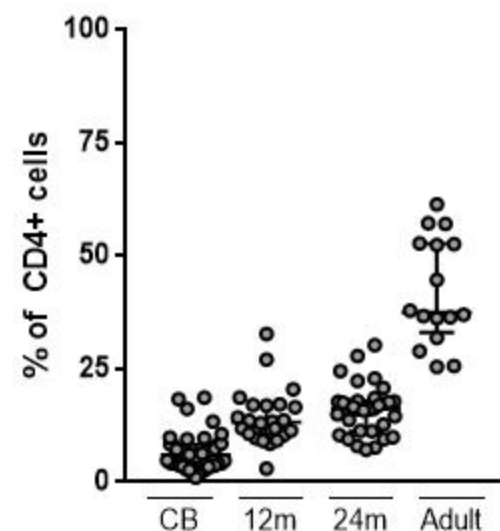
Proportion CD4+CD45RA+
(naïve) cells



n (samples) = 44 24 33 16

CD4CD45RA+ (naïve)		CB	12m	24m	Adult
	CB		***	***	***
	12m	***		ns	***
	24m	***	ns		***
	Adult	***	***	***	

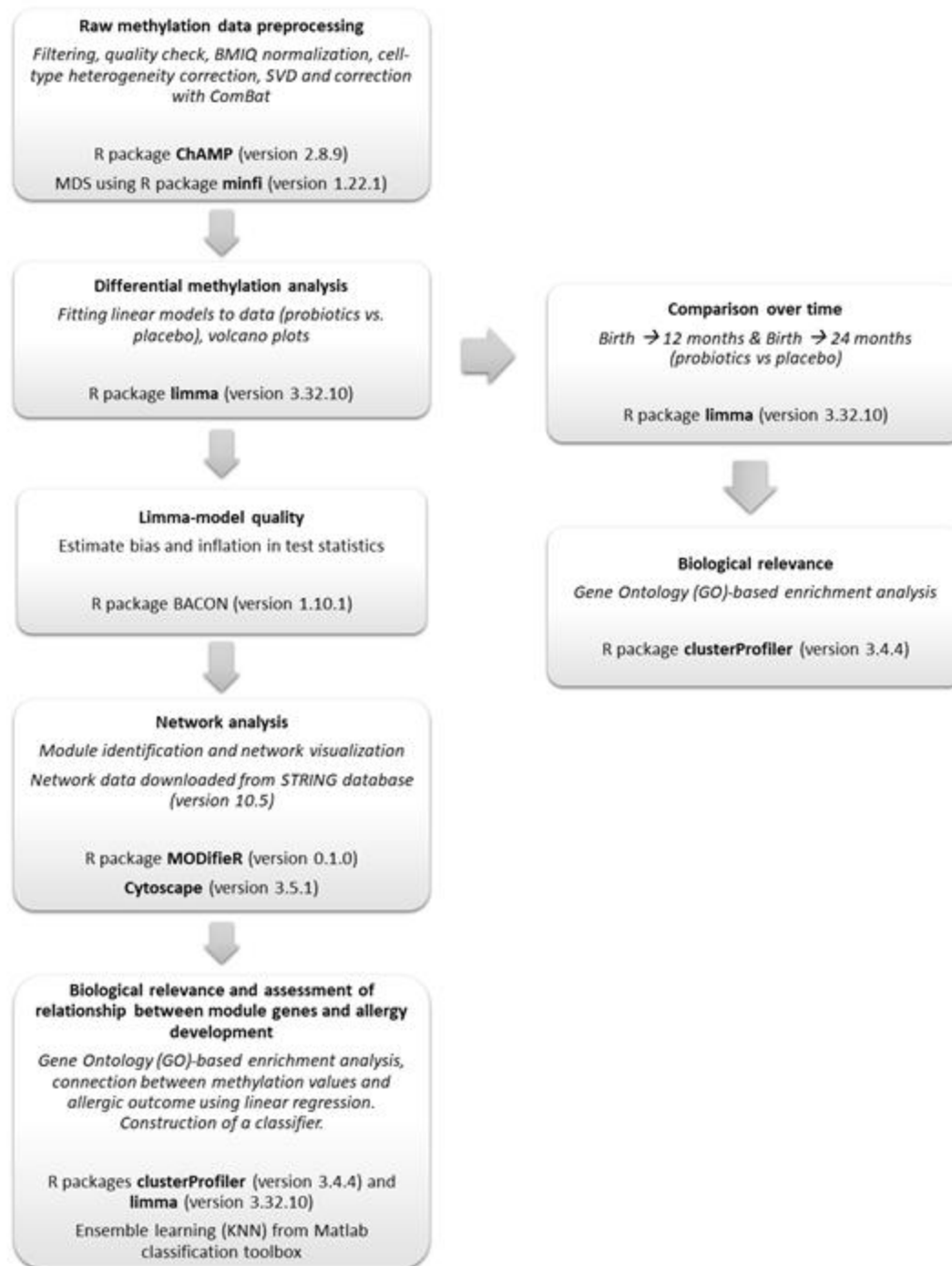
Proportion CD4+CD45RA-
(memory) cells



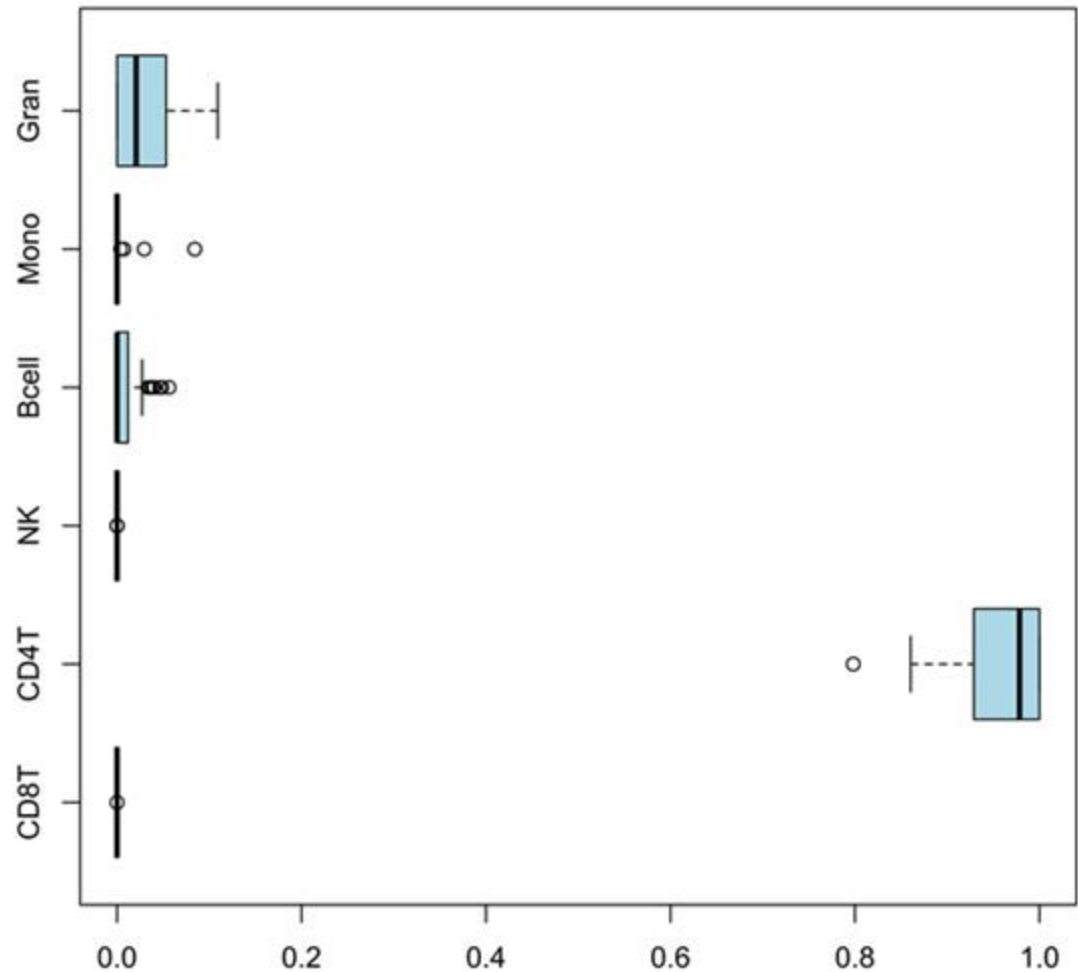
n (samples) = 44 24 33 16

CD4CD45RA- (memory)		CB	12m	24m	Adult
	CB		***	***	***
	12m	***		ns	***
	24m	***	ns		***
	Adult	***	***	***	

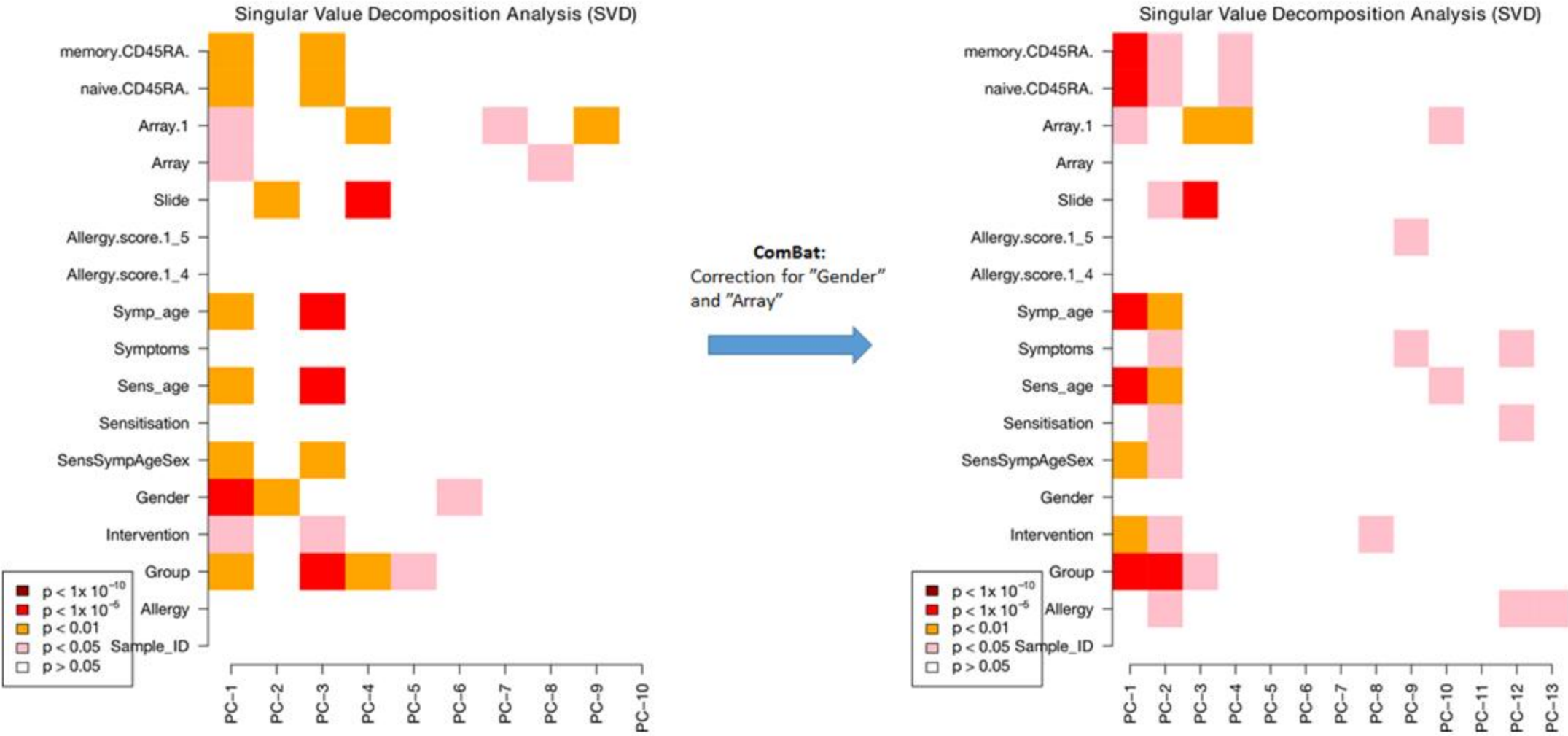
Supplementary Figure 8.



Supplementary Figure 9.



Supplementary Figure 10.



Supplementary table I. Summary of extracted differentially methylated probes (DMPs) for nine different group comparisons and different stringency levels of significance upon FDR correction for multiple comparisons. DMPs were defined as such upon fulfilling an FDR-significance level along with a mean log2 methylation difference (MMD) of $\pm 5\%$.

	FDR<0.1 MMD > $ \pm 0.05 $	FDR<0.05 MMD > $ \pm 0.05 $	FDR<0.01 MMD > $ \pm 0.05 $
1. <i>Probiotics vs Placebo (CB)</i>	1629	277	8
2. <i>Probiotics vs Placebo (12m)</i>	0	0	0
3. <i>Probiotics vs Placebo (24m)</i>	1	0	0
4. <i>CB vs 12m (Probiotics)</i>	9970	7847	4461
5. <i>CB vs 24m (Probiotics)</i>	22165	19848	15397
6. <i>12m vs 24m (Probiotics)</i>	21	12	0
7. <i>CB vs 12m (Placebo)</i>	15838	13257	8782
8. <i>CB vs 24m (Placebo)</i>	18667	16769	12959
9. <i>12m vs 24m (Placebo)</i>	19	16	9

CB – cord blood, DMP – differentially methylated probe, FDR – false discovery rate,
MMD – mean log2 methylation difference.

Supplementary Table XV. Descriptive data of the children in the T helper subset locus-specific DNA methylation and Illumina 450K analyses, stratified by age and intervention group per method. No significant differences were evident when comparing intervention allotment in the respective methods at any age to the same in the large study.

	T helper subset locus-specific DNA methylation analysis % (n/N)						Illumina 450K array % (n/N)					
	CB		12m		24m		CB		12m		24m	
	Placebo	Probiotics	Placebo	Probiotics	Placebo	Probiotics	Placebo	Probiotics	Placebo	Probiotics	Placebo	Probiotics
Intervention	56 (18/32)	44 (14/32)	39 (9/23)	61 (14/23)	43 (15/35)	57 (20/35)	53 (10/19)	47 (9/19)	60 (6/10)	40 (4/10)	58 (11/19)	42 (8/19)
Boys	60 (6/10)	40 (4/10)	57 (4/7)	43 (3/7)	36 (5/14)	64 (9/14)	50 (4/8)	50 (4/8)	100 (2/2)	0 (0/2)	40 (2/5)	60 (3/5)
First born	53 (9/17)	47 (8/17)	42 (5/12)	58 (7/12)	41 (7/17)	59 (10/17)	44 (4/9)	56 (5/9)	60 (3/5)	40 (2/5)	67 (6/9)	33 (3/9)
Parental smoking birth	100 (2/2)	0 (0/2)	0 (0/1)	100 (1/1)	67 (2/3)	33 (1/3)	0 (0/0)	0 (0/0)	0 (0/0)	0 (0/0)	0 (0/1)	100 (1/1)
Furred pets	100 (2/2)	0 (0/2)	60 (3/5)	40 (2/5)	40 (2/5)	60 (3/5)	100 (1/1)	0 (0/1)	75 (3/4)	25 (1/4)	50 (2/4)	50 (2/4)
Maternal atopy	60 (15/25)	40 (10/25)	45 (9/20)	55 (11/20)	38 (12/32)	62 (20/32)	56 (9/16)	44 (7/16)	60 (6/10)	40 (4/10)	53 (9/17)	47 (8/17)
Breastfeeding 3m	65 (11/17)	35 (6/17)	46 (6/13)	54 (7/13)	50 (12/24)	50 (12/24)	67 (6/9)	33 (3/9)	71 (5/7)	29 (2/7)	67 (10/15)	33 (5/15)
Breastfeeding 6m	63 (12/19)	37 (7/19)	42 (8/19)	58 (11/19)	44 (14/32)	56 (18/32)	60 (6/10)	40 (4/10)	63 (5/8)	37 (3/8)	61 (11/18)	39 (7/18)
Antibiotics 0-12m	50 (2/4)	50 (2/4)	33 (2/6)	67 (4/6)	25 (2/8)	75 (6/8)	0 (0/2)	100 (2/2)	67 (2/3)	33 (1/3)	40 (2/5)	60 (3/5)
Antibiotics 12-24m	62 (8/13)	38 (5/13)	42 (5/12)	58 (7/12)	33 (4/12)	67 (8/12)	57 (4/7)	43 (3/7)	40 (2/5)	60 (3/5)	40 (2/5)	60 (3/5)
Day-care 0-12m	0 (0/0)	0 (0/0)	0 (0/0)	0 (0/0)	100 (1/1)	0 (0/1)	0 (0/0)	0 (0/0)	0 (0/0)	0 (0/0)	0 (0/0)	0 (0/0)
Day-care 12-24m	62 (13/21)	38 (8/21)	38 (8/21)	62 (13/21)	42 (13/31)	58 (18/31)	55 (6/11)	45 (5/11)	56 (5/9)	44 (4/9)	56 (9/16)	44 (7/16)

Supplementary Table XVI. Adult allergic individuals with age- and sex-matched non-allergic adults, included for comparison with T helper subset locus-specific DNA methylation analysis data from infants.

Individual	Age	Sex	Allergens	Medication	Symptoms
ALLERGIC SUBJECTS					
A1	34	F	Birch, grass, cat, dog, mug wort	Antihistamines	AD, ARC
A2	52	F	Mould, grass	Antihistamines, inhaled corticosteroids	AB, ARC
A3	23	M	Birch, timothy-grass, cat, dog, hazel nut, almond, peanut, stone fruits	Antihistamines	ARC, FA
A4	24	F	Egg, milk protein, birch, mites, timothy-grass, horse, cat, dog, almond	Antihistamines, cortisol cream	AB, AD, ARC, FA, U
A5	27	F	Peanut, dog, cat, horse, birch, timothy-grass	Antihistamines	AD, ARC, FA, U
A6	25	F	Dog, horse, cat, hazel nut, almond, peanut, green peas	Inhaled corticosteroids	AB, AD, ARC, FA
A7	23	F	Birch, grass, dairy products, eggs, nuts, almonds	Inhaled corticosteroids, nasal spray (glucocorticoid), antihistamines	AB, ARC, FA, U
A8	19	F	Pollen, cat, guinea pig	Antihistamines, anti-leukotrienes, inhaled corticosteroids, anti-IgE, beta-2-receptor agonist	AB, ARC, FA, U
NON-ALLERGIC CONTROL SUBJECTS					
NA1	38	F	-	-	-
NA2	49	F	-	-	-
NA3	26	M	-	-	-
NA4	24	F	-	-	-
NA5	28	F	-	-	-
NA6	25	F	-	-	-
NA7	26	F	-	-	-
NA8	19	F	-	-	-

AB – asthma bronchiale, AD – atopic dermatitis, ARC – allergic rhinoconjunctivitis, FA – food allergy, U - urticaria

Supplementary table XVII. Table displaying mean proportions (in %) of different immune cell populations in the analysed samples in the Illumina 450K study at birth (cord blood), 12 and 24 months, respectively. Calculations were made according to Houseman *et al.* 2012.

Cell type	Mean cell proportion (%)					
	Cord blood		12 months		24 months	
	Prob	Plac	Prob	Plac	Prob	Plac
CD8 T cell	5.47e-19	0	0	0	0	0
CD4 T cell	0.95	0.92	1	0.99	0.99	0.98
NK cell	4.51e-18	1.56e-18	0	1.18e-18	3.34e-18	1.18e-18
B cell	0.01	0.01	0	0.002	0.005	0.008
Monocyte	5.72e-18	0.0004	2.11e-18	0	0.004	2.32e-19
Granulocyte	0.04	0.07	0	0.003	0.005	0.009

1 **Online supplementary material**

2 **Supplementary figure legends**

3 **Supplementary Figure 1.** A Q-Q-plot for visualisation of obtained p-values of the linear
4 limma model at birth (probiotics vs. placebo, with cell type as co-variate) against a normal
5 distribution. The genomic p-value inflation factor was also calculated using the R package
6 BACON, indicating a slight over-significance of the data (estimated inflation factor= 1.2,
7 estimated bias=-0.003).

8 **Supplementary Figure 2.** Separate within-group comparisons of DNA methylation patterns
9 over time and subsequent pathway enrichment. A. shows volcano plots of differential
10 methylation in each probe (-log10 FDR-adjusted p-values plotted against mean log2
11 methylation difference (MMD)) for over time comparisons; probiotics (left) and placebo (right)
12 comparisons are depicted vertically, CB vs. 12m (top) and CB vs. 24m (bottom) horizontally.
13 Differential methylation analysis included fitting a linear model to data and FDR-correction for
14 multiple testing. The horizontal dashed line represents a cut-off of $FDR < 0.01$, the vertical
15 dashed lines correspond to $MMD < -0.05$ and > 0.05 , respectively. Red dots in the graph
16 indicate differentially methylated probes fulfilling both the FDR- and MMD-criteria mentioned
17 above. The resulting DMPs were mapped to corresponding genes, and further analysed for
18 GO-pathway enrichment. In B. a dot plot displays the enrichment results based on genes
19 containing significant DMPs in over-time comparisons between birth and 12 months. The x-
20 axis represents different gene sets originating from the placebo group (PlacOnly) or genes
21 that were overlapping between the intervention groups (Intersect). Dot size corresponds to
22 gene ratio, dot color represents adjusted enrichment p-values. No enrichment was obtained
23 from the genes of the over-time comparisons in the probiotic group. DMP- differentially
24 methylated probe, FDR – False Discovery Rate, GO – gene ontology, MMD – mean log2
25 methylation difference.

26 **Supplementary Figure 3.** Venn diagrams and dot plots for DMPs in the separate within-
27 group comparisons over time. In A. and B. the Venn diagrams illustrate the distribution of the
28 differentially methylated probes in each of the intervention groups, as well as overlapping
29 DMPs for CB vs. 12m (A) and CB vs. 24m (B). In C. and D. dot plots illustrate the distribution
30 of the overlapping DMPs in the over-time comparisons for CB vs. 12m (C) and 24m (D),
31 respectively. The direction of the overlapping genes are consistently in the same direction,
32 regardless of intervention group. CB – cord blood, DMP – differentially methylated probe.

33 **Supplementary Figure 4.** A dot plot of pathway enrichment from the within-group
34 comparisons from birth to 24 months. Significant DMPs ($FDR < 0.01 + MMD < -0.05$ or $>$

0.05) from the birth to 24 months comparison within each group and in the overlapping intersect were mapped to corresponding genes and analysed for GO-pathway enrichment. The x-axis represents different gene sets originating from the placebo group (PlacOnly), probiotic group (ProbiOnly) or genes that were overlapping between the intervention groups (Intersect). Dot size corresponds to gene ratio, dot colour represents adjusted enrichment p-values. No enrichment was obtained from hypermethylated DMPs in genes of the over-time comparisons in the probiotic group. DMP– differentially methylated probe, FDR – False Discovery Rate, GO – gene ontology, MMD – mean log2 methylation difference.

Supplementary Figure 5. Graphical illustration of the distribution of differential methylation values among DMPs in each gene from the between group comparison (probiotics vs. placebo) at birth. The X-axis represents genes sorted in order of mean differential methylation as log fold given by the Y-axis change. The X-axis' gene annotations have been omitted for visualization reasons. The DMPs were found in a total of 1015 different genes, of which 665 genes (65.5%) contained probes that were only hypomethylated, 331 genes (32.6%) contained probes that were only hypermethylated, and 19 genes (1.9%) exhibited DNA methylation patterns of both types. DMP – differentially methylated probe.

Supplementary Figure 6. Gating strategy for analysis of CD45RA+/- expression in sorted CD4+ cells. FSC and SSC were used to determine the lymphocyte population along with DAPI to only include live cells. CD4 PerCP and CD45RA FITC expression further gated the cells to distinguish memory (CD45RA-) and naïve (CD45RA+) populations. FSC – forward scatter, SSC – side scatter.

Supplementary Figure 7. Flow cytometry data of sorted CD4+ cells in cord blood at birth and in peripheral blood at 12 and 24 months of age and in adults. A. depicts naïve (CD4+CD45RA+) cells and B. memory (CD4+CD45RA-) cells. A One-way ANOVA with Tukey's correction for multiple testing was employed for the analysis. Adjusted p-values are reported. *** $p < 0.001$

Supplementary Figure 8. Pipeline for the workflow of the DNA methylation analyses from the Illumina 450K platform. Each step in the workflow gives a brief description of the analyses included, as well as the specific R packages used. See Appendix 1 for more detailed descriptions.

Supplementary Figure 9. Plot illustrating proportions of different immune cell populations in the analysed samples utilised in the Illumina 450K study, as retrieved by cell proportion analyses according to Houseman *et al.* 2012.

68 **Supplementary Figure 10.** Results of Singular Value Decomposition analyses before and
69 after ComBat correction for batch effects and sex (variable here named gender).

70

71 ***Supplementary table legends***

72 **Supplementary table I.** Summary of extracted differentially methylated probes (DMPs) for
73 nine different group comparisons and different stringency levels of significance upon FDR
74 correction for multiple comparisons. DMPs were defined as such upon fulfilling an FDR-
75 significance level along with a mean log2 methylation difference (MMD) of $\pm 5\%$.

76 **Supplementary table II.** Table of differentially methylated probes (DMPs), defined as mean
77 log2 methylation difference (MMD) $> \pm 5\%$ and $FDR < 0.1$, extracted for the comparison of
78 probiotics vs. placebo at birth (in cord blood).

79 **Supplementary table III.** Table of differentially methylated probes (DMPs), defined as mean
80 log2 methylation difference (MMD) $> \pm 5\%$ and $FDR < 0.1$, extracted for the comparison of
81 probiotics vs. placebo at 12 months of age.

82 **Supplementary table IV.** Table of differentially methylated probes (DMPs), defined as mean
83 log2 methylation difference (MMD) $> \pm 5\%$ and $FDR < 0.1$, extracted for the comparison of
84 probiotics vs. placebo at 24 months of age.

85 **Supplementary table V.** Table of differentially methylated probes (DMPs), defined as mean
86 log2 methylation difference (MMD) $> \pm 5\%$ and $FDR < 0.1$, extracted for the comparison
87 between the time of birth and 12 months within the probiotic group.

88 **Supplementary table VI.** Table of differentially methylated probes (DMPs), defined as mean
89 log2 methylation difference (MMD) $> \pm 5\%$ and $FDR < 0.1$, extracted for the comparison
90 between the time of birth and 24 months within the probiotic group.

91 **Supplementary table VII.** Table of differentially methylated probes (DMPs), defined as
92 mean log2 methylation difference (MMD) $> \pm 5\%$ and $FDR < 0.1$, extracted for the comparison
93 between 12 and 24 months of age within the probiotic group.

94 **Supplementary table VIII.** Table of differentially methylated probes (DMPs), defined as
95 mean log2 methylation difference (MMD) $> \pm 5\%$ and $FDR < 0.1$, extracted for the comparison
96 between the time of birth and 12 months within the placebo group.

97 **Supplementary table IX.** Table of differentially methylated probes (DMPs), defined as mean
98 log2 methylation difference (MMD) $> \pm 5\%$ and FDR <0.1 , extracted for the comparison
99 between the time of birth and 24 months within the placebo group.

100 **Supplementary table X.** Table of differentially methylated probes (DMPs), defined as mean
101 log2 methylation difference (MMD) $> \pm 5\%$ and FDR <0.1 , extracted for the comparison
102 between 12 and 24 months of age within the placebo group.

103 **Supplementary table XI.** A table of GO-terms from enrichment analyses of the 1015 genes
104 which originated from the 1629 significant DMPs found in the probiotics vs. placebo
105 comparison at the time of birth.

106 **Supplementary table XII.** List of all 91 genes provided by the consensus module in the
107 network analyses, given in both gene symbols and Ensembl IDs. For each gene a color has
108 been assigned to describe differential methylation in both the probiotics vs. placebo and the
109 allergy vs. non-allergy analysis, respectively. Blue corresponds to a presence of only hypo-
110 methylated DMPs, red corresponds to a presence of only hyper-methylated DMPs, and a
111 color-marking of both blue and red is assigned to genes containing both hypo- and hyper-
112 methylated DMPs in the respective analysis. The source of the gene data is also stated,
113 either as originating from differential methylation analyses, or from the network analyses. The
114 number of connections in the network is also stated. * indicates that the gene is related to
115 allergy development. See also Supplementary table XIII for further information on the allergy
116 related genes.

117 **Supplementary table XIII.** A list of genes from the network analyses with at least 5% mean
118 methylation difference and a high number of connections within the network, which revealed
119 previous associations to allergy outcomes in current literature.

120 **Supplementary table XIV.** A comprehensive list of the included study subjects, including
121 information about atopic manifestations at 2 and 7 years of age, intervention groups, allergic
122 status as well as sample availability for the T helper subset locus-specific DNA methylation
123 and Illumina analyses, respectively. For the classifier analyses of the Illumina samples, in
124 total 41 samples were included of which 17 individuals were healthy (nProb = 5, nPlac = 12)
125 and 24 (nProb = 12, nPlac = 12) individuals developed allergic manifestations (cumulatively
126 up to 7 years of age). AD – atopic dermatitis, ARC – allergic rhinoconjunctivitis, CB – cord
127 blood, GI – gastrointestinal, SPT – skin prick test.

128 **Supplementary table XV.** Descriptive data of the children in the T helper subset locus-
129 specific DNA methylation and Illumina 450K analyses, stratified by age and intervention

group per method. No significant differences were evident when comparing intervention allotment in the respective methods at any age to the same in the large study.

Supplementary table XVI. A descriptive table of adult (allergic and non-allergic, age- and sex-matched) samples included in the study for comparison with infant samples in the T helper subset locus-specific DNA methylation analysis.

Supplementary table XVII. Table displaying mean proportions (in %) of different immune cell populations in the analysed samples in the Illumina 450K study at birth (cord blood), 12 and 24 months, respectively. Calculations were made according to Houseman *et al.* 2012.

Detailed information on experimental and statistical procedures

Study design

Blood cell samples were included from cord blood, and at 12 months and 24 months of age from infants in a double-blind, randomised, placebo-controlled probiotic trial, in which 188 infants completed the study (ClinicalTrials.gov, NCT01285830).¹ The participants in the present study were representative of the main study for all but four variables: gender at birth and 12 months of age, maternal atopy at 24 months in the T helper subset locus-specific DNA methylation analyses, and furred pet ownership at 12 months in the Illumina 450K analyses (Table I, Supplementary Table XV for stratification on intervention groups). Informed consent was obtained from both parents before inclusion. Infants included in the study had a family history of allergic disease, *i.e.* at least one family member had eczema, asthma, gastrointestinal allergy, allergic urticaria or allergic rhinoconjunctivitis. The pregnant women either received *Lactobacillus reuteri* ATCC 55730 or placebo daily from week 36 until delivery, and their infants started on day 1-3 with the same product daily until 12 months of age (Figure 1). The probiotic preparation consisted of freeze-dried *L. reuteri* suspended in coconut and peanut oil containing cryoprotective components, while the placebo preparation only comprised the vehicle. Daily intake corresponded to 1×10^8 colony-forming units. The children were monitored by research nurses and follow-ups were performed by a paediatrician at 2 and 7 years of age. Skin-prick tests were performed at 6, 12 and 24 months and 7 years of age. Allergic diseases included eczema, allergic rhinoconjunctivitis, urticaria, asthma and gastrointestinal allergy. For detailed information on definitions of allergic disease, we refer to previously published papers on the probiotic trial.^{1,2} Available blood samples were selected for analysis with the T helper subset locus-specific DNA methylation analysis (n=91) and/or Illumina 450K analysis (n=48, see Supplementary Table XIV, Figure 1 and T helper subset locus-specific DNA methylation analysis/Illumina 450K

sections below). Additionally, T helper subset locus-specific DNA methylation analysis results from adults (age- and sex-matched, mean age=28.9 years, n Allergic=8, n Healthy=8) were included for a comparison of immune maturation later in life (Supplementary Table XVI). The study was approved by the Regional Ethics Committee for Human Research in Linköping, Sweden (Dnr: 99323 and M171-07).

Isolation of CD4+ cells from CBMCs/PBMCs

From the heparinized cord blood and whole-blood samples, cord blood mononuclear cells (CBMCs) and peripheral blood mononuclear cells (PBMCs) were isolated as previously described³ and stored in liquid nitrogen until further use. Then, CBMCs/PBMCs were thawed in a 37°C water-bath and put in RPMI (Gibco Life Technologies, Paisley, UK) + 10% FCS (Sigma-Aldrich, Ayrshire, United Kingdom) and centrifuged for 10 min (400 g, 4°C). The pellet was resuspended in 3 ml HBSS (Gibco Life Technologies, Paisley, UK) and filtered through a MACS Pre-separation filter (Miltenyi Biotec, Bergisch-Gladbach, Germany). Cells were counted and centrifuged for 10 min. The supernatant was removed, and the pellet was resuspended in 80 µl MACS-buffer per 10 million PBMCs. 20 µl CD4+ magnetic Microbeads (Miltenyi Biotec) per 10 million PBMCs was added, the solution was incubated for 15 min at 8°C, and placed on a MACS Separation Column (Miltenyi Biotec). The collected CD4+ cells were centrifuged, the supernatant removed and the pellet resuspended in PBS (Medicago AB, Uppsala, Sweden) to a concentration of 10×10^6 cells/ml. A number of 0.2×10^6 cells were used for flow cytometry analyses, using CD4 PerCP (SK3), CD45RA FITC (L48), CD45RO PeCy7 (UCHL-1), with a complementary isotype tube including CD4 PerCP, Mouse $\gamma 1$ FITC (X40) and Mouse IgG2a PeCy7 (G155-178) (all antibodies: Becton-Dickinson, San José, CA, USA). DAPI (Becton-Dickinson) was also added to each tube, in order to along with FSC and SSC determine the lymphocyte population and only include live cells. For gating strategy see Supplementary Figure 6 and for results see Supplementary Figure 7. The remaining cells repeated the centrifugation cycle. The supernatant was removed and the cells were resuspended in freezing medium and stored in -20°C until DNA extraction. From these analyses, missing data originate from one cord blood sample and one 24 months sample.

DNA extraction

The CD4+ cells were thawed on ice and centrifuged for 5 min, 13×10^5 g at room temperature (RT), and then washed in ice cold PBS (Medicago AB, Uppsala, Sweden) and centrifuged for another 5 min at RT. The supernatant was then removed, and the pellet was resuspended in

198 100 µl digestion buffer and 2 µl 0.1 mg/ml Proteinase K (Zymo Research, Irvine, CA, USA).
199 Cells were vortexed and incubated overnight at 50°C while shaking. The next day, 100 µl
200 phenol/chloroform/isoamyl alcohol (25:24:1 saturated with 10 mM Tris, pH 8.0, 1 mM EDTA,
201 Sigma-Aldrich, St Louis, Missouri, USA) was added, and the samples were vortexed and
202 centrifuged for 10 min at 4°C. The aqueous top layer was transferred to a new Eppendorf
203 tube and the volume of the sample was measured using a pipette. 3M sodium acetate
204 (Merck, Darmstadt, Germany) was added in the ratio of 0.5:1 and 100% ethanol was added
205 in the ratio of 2:1 to the sample. Then, a centrifugation at RT for 10 min followed, and the
206 supernatant was discarded. The samples were let to air dry in RT until the liquid had
207 evaporated, and were then dissolved in 30 µl H₂O and shaken at 900 rpm and 50°C
208 overnight.

209

210 **T helper subset locus-specific DNA methylation analysis**

211 In order to study the epigenetic effects on polarisation into specific T helper cell subsets,
212 specific CpG dinucleotide residues located at regulatory sites upstream of the respective
213 promoters of the *IFNG* (position -4229 TSS⁴), *IL13* (position -46 TSS^{5,6}), *IL17* (position -122
214 TSS⁶) and *FOXP3* (position -77 TSS⁷) loci were investigated regarding their methylation
215 status, in accordance with previous studies by Janson et al.⁶ This procedure, known as
216 epigenetic immune lineage analysis, comprises a number of sequential experiments. Initially,
217 CD4+ T cells were isolated from cord blood mononuclear cells (CBMCs) or peripheral blood
218 mononuclear cells (PBMCs). Thereafter, DNA was extracted from these cells, and
219 subsequently bisulphite converted. This reaction results in the conversion of unmethylated
220 cytosine residues into uracil, leaving methylated cytosine unaltered. In the following locus-
221 specific PCR reaction, where each locus is amplified separately, the uracil residues will be
222 translated into thymine residues. Thereafter, in the methylation sensitive single nucleotide
223 polymorphism primer extension (Ms-SNuPE/SNaPshot) reaction, unlabelled primers bind
224 adjacent to the chosen CpG dinucleotides and are extended with fluorescently labelled
225 ddNTPs.⁸ Due to differing lengths of the primers used for the different loci, the methylation
226 status, *i.e.* the corresponding incorporated fluorescently labelled ddNTPs of each locus, may
227 be identified by means of separation using capillary electrophoresis. The four investigated
228 loci are not included in the Illumina 450K array. For these analyses, a total of 91 blood
229 samples were included from birth (cord blood, n=32), 12 months (n=24) and 24 months
230 (n=35) (see Figure 1, Supplementary Table XIV). All four loci were successfully detected in
231 all samples, with the exception of the following samples. *IL13* in four samples (CB=1,

232 24m=3), *IL17* in five samples (CB=1, 12m=4), and for *IFNG* and *FOXP3* one sample each at
233 24m.

234

235 **Bisulphite conversion**

236 For the bisulphite conversion, an EZ DNA Methylation-Gold kit (Zymo Research, Tustin, CA,
237 USA) was used according to the manufacturer's protocol, except for incubating at 64°C for
238 3.5 h (instead of 2.5 h) and eluting the DNA in 12 µl M-Elution Buffer. In every bisulphite
239 conversion reaction, 350 ng of genomic DNA was used.

240

241 **Locus-specific PCR**

242 In order to investigate the methylation status of the four loci in promoter regions upstream of
243 the genes of interest, *i.e.* *IL13* (position -46 TSS)⁵, *IL17* (position -122 TSS)⁶, *FOXP3* (position
244 -77 TSS)⁷ and *IFNG* (position -4229 TSS)⁴, the loci were amplified by means of a locus-
245 specific PCR. To a PCR tube, 6.25 µl of a ready-to-use PCR buffer (GoTaq HotStart
246 Colorless Mastermix, Promega, Madison, WI, USA), 1 µl of forward and 1 µl of reverse
247 primers for the respective loci (see Primer sequence specification below), 1.25 µl of BSA, 1
248 µl of RNase-free H₂O and 2 µl of bisulphite-converted DNA was added. The samples were
249 processed at 49.0°C (*IL-17* locus) and 53.1°C (the other three loci) in the PCR machine (Artik
250 Thermal Cycler, Thermo Scientific, Waltham, MA, USA). To clean up the PCR products from
251 unincorporated primers and unwanted single-stranded DNA, 2 µl of the Illustra ExoProStar 1-
252 step reagent (GE Healthcare, Buckinghamshire, UK) was added to the PCR tube, which was
253 then processed in the PCR machine (see PCR program specifications below for details).

254

255 **Ms-SNuPE/SNaPshot and CIP post extension**

256 To examine the methylation status of the amplified loci, a Methylation sensitive Single
257 Nucleotide Primer Extension (Ms-SNuPE) was performed, using the SNaPshot multiplex kit
258 (Applied Biosystems, Warrington, United Kingdom). The reaction was prepared on ice, and
259 to every reaction 5 µl Snapshot multiplex ready reaction mix and 1 µl of pooled Snapshot
260 primers (2 µM, see see Primer sequence specification below for more information) was
261 added to 4 µl of PCR product (1 µl from each locus). The reaction was then run through the
262 Snapshot PCR programme (see PCR program specifications below). A calf-intestinal
263 phosphatase (CIP, 1000 U/ml, 1U/sample, New England Biolabs, Ipswich, MA, USA) post-

264 extension reaction was then performed, in order to remove the remaining reactants from the
265 Snapshot reaction. This reaction was run through the CIP programme in the PCR machine
266 (see PCR program specifications below).

267

268 **Capillary electrophoresis**

269 The samples were analysed in an ABI Prism 3500 Genetic Analyzer (Applied Biosystems,
270 Carlsbad, CA, USA). 12.5 µl HiDi formamide (Applied Biosystems) and 0.5 µl GeneScan 120
271 LIZ (Applied Biosystems) was added to 2 µl Snapshot product, and the samples were then
272 added to a Microamp optical 96 well plate (Applied Biosystems). A 2 min centrifugation at
273 400 rpm (RT) followed. The plate was then incubated for 2 min at 95°C, and cooled down on
274 ice, prior to analysis in the Genetic Analyzer 3500. The results were analysed using the
275 Gene Mapper software (version 4.1, Applied Biosystems).

276

277 **Illumina 450K (Infinium Human Methylation 450K bead chip)**

278 For the study of global DNA methylation patterns on >485 000 sites, Illumina's Infinium
279 Human Methylation 450K bead chip was utilised. Genomic DNA was extracted from isolated
280 CD4+ T helper cells (as previously described, Figure 1) from 48 blood samples retrieved at
281 birth (cord blood, n=19), at 12 months (n=10) and at 24 months (n=19) (see, Supplementary
282 Table XIV). The samples were sent to the SNP&SEQ-technology platform at SciLifeLab
283 (Uppsala University, Uppsala, Sweden) for the performance of Illumina's Infinium Human
284 Methylation 450K bead chip array preceded by on site bisulphite conversion. The amount of
285 DNA used for each reaction was in the range from 100-250 ng.

286

287 **Statistics**

288 **Flow cytometry and T helper subset locus-specific DNA methylation analysis**

289 As flow cytometry data were normally distributed, a One-way ANOVA with Tukey's correction
290 for multiple testing was used for the statistical analysis. The data from the T helper subset
291 locus-specific DNA methylation analyses were mostly not normally distributed, motivating the
292 use of the non-parametric Kruskal-Wallis test, followed by the two-stage linear step-up false
293 discovery rate (FDR) procedure of Benjamini, Krieger and Yekutieli.⁹ GraphPad Prism 7
294 (Version 7.04) was used for the analyses.

295

296 **Illumina 450K**

297 For a summary of workflow, see the pipeline in Supplementary Figure 8. In short, the DNA
298 methylation data were pre-processed prior to initial multidimensional scale (MDS) analyses
299 to elucidate general differences in DNA methylation patterns. Thereafter, differential
300 methylation analyses of comparisons between the intervention groups were performed, as
301 well as over-time comparisons within each intervention group using linear models
302 (Supplementary Table I). For all comparisons, quality control measures and bioinformatic
303 analyses to elucidate biological relevance were pursued. Network analyses were performed
304 after module identification for analyses between intervention groups and further related to the
305 development of allergic disease up to 7 years of age, also by building a classifier based on
306 the network analyses.

307

308 **Pre-processing and quality control**

309 Preparation and pre-processing of the raw DNA methylation data were performed using the
310 Chip Analysis Methylation Pipeline (ChAMP)¹⁰ R package (version 2.8.9). This included
311 filtering probes that failed a detection p-value < 0.01 and removing samples with a high
312 (>5%) fraction of failing probes, removal of non-CpG probes and filtering of multi-hit probes.
313 In addition, SNP-related probes that may contribute to genetic rather than epigenetic
314 observations were also removed. Probes situated on the X and Y chromosomes were
315 retained. Out of the initial 485 577 probes found on the Illumina HM450 BeadArray a total of
316 420 619 probes were left after the pre-processing steps. No samples exhibited a fraction of
317 failed probes above 5%, and so all samples were retained. The data was normalised using
318 the Beta-Mixture Quantile dilation (BMIQ) normalisation method.¹¹ Data quality checks were
319 performed prior to and after the normalisation step.

320 Cell heterogeneity in the data was evaluated by applying a reference-based method based
321 on the work of Houseman *et al.* 2012.¹² This was done in order to investigate whether the
322 isolation of CD4⁺ T cells had produced sufficiently pure samples or if correction had to be
323 performed. The analysis revealed that, in the collective sample setup as well as separately
324 per age, the average proportion of CD4⁺ T cells in samples was more than 95%
325 (Supplementary Figure 9 and Supplementary Table XVII), and therefore correction was not
326 considered necessary. Possible contamination of the cord blood samples by maternal cells
327 was investigated according to the method suggested by Morin *et al.* 2017.¹³ The methylation
328 values of 10 specific CpG sites upon BMIQ-normalisation were related to defined threshold

values. A sample is considered as contaminated if at least 5 out of 10 CpGs are above the defined threshold values. None of the included 19 cord blood samples had values above the threshold values for any of the CpGs, suggesting no contamination of the samples.

Single value decomposition (SVD) was used to identify most significant components of variation in the data and correction for unwanted sources of variation was done using the ComBat normalisation method.¹⁴ Based on the SVD results (Supplementary Figure 10) ComBat-correction was performed for variation in data due to batch effects and sex. Multidimensional scaling (MDS) was performed separately at the different time points using the R package minfi (version 1.22.1)¹⁵ on all available samples after initial filtering, BMIQ-normalisation and cell type correction of raw data, hence upon pre-processing prior to the performance of high-level analyses. This was done in order to elucidate whether there were any inherent differences of the data to begin with. Upon the statistical analyses, the MDS results were graphically represented by assigning the samples to their respective intervention group belonging

Differential methylation analysis

Differentially methylated probes (DMPs) were identified by fitting a linear model to the data using the limma¹⁶ R package (version 3.32.10), obtaining moderated t-statistics, log₂ Fold Change (logFC) values and associated p-values for each probe position. P-values were adjusted for multiple testing using Benjamini-Hochberg False Discovery Rate (FDR) correction. The logFC values are defined as the difference between the probiotic and the placebo group in mean logarithmic methylation values. It may be interpreted as a change in the log odds of methylation values, and hence indirectly represents a change in average beta values between the intervention groups. We from here on refer to this as the mean Log₂ methylation difference (MMD). DMPs were defined as probes exhibiting both an MMD of $> \pm 5\%$ along with an FDR-corrected p -value < 0.1 . The quality properties for the cord blood (probiotics vs. placebo) limma model were assessed using the QCEWAS R package.¹⁷ A normal quantile-quantile (Q-Q) plot was constructed, plotting the distribution of obtained p-values from the limma model against a normal distribution. As a means to assess the inflation of the obtained test-statistics, calculations of the genomic inflation factor were performed. While the commonly used genomic inflation factor has been found to often overestimate the inflation in EWAS and TWAS, the recently developed R-package BACON¹⁸ was used to estimate inflation and bias in our test-statistics. Furthermore, in order to assure statistical robustness of the differentially methylated sites identified, bootstrap and permutation P-value calculations were performed using MATLABs (version 9.6) mattest

function using 100 000 re-shufflings. In order to account for variation due to differences in T cell subtypes (Supplementary figure 5), the linear model included flow cytometry data on proportions of naïve CD4⁺ T cells and memory CD4⁺ T cells as covariates (Supplementary figure 7). Nine comparisons were defined based on treatment (probiotics or placebo) and time of sampling (CB, 12m or 24m) (Supplementary table I). Comparisons of the development of atopic manifestations (sensitisation and/or symptoms cumulatively up to 7 years of age) were also performed separately at the different time points. Results obtained when comparing time points for each of the treatments separately yielded a large number of DMPs and hence a stricter threshold of FDR<0.01 was used for significance. Significant DMPs were mapped to genomic regions and genes containing at least one DMP were extracted. The biological relevance of these findings was investigated using online tools of the Database for Annotation, Visualization and Integrated Discovery (DAVID, version 6.8).^{19,20}

Network analysis

In order to further elucidate the potential functional role of the methylation changes, DMP-containing genes were placed within and analysed from a network context by elaborating a disease module. Briefly, a disease module is defined as a set of disease-related genes that co-localise in the protein-protein interaction (PPI) network and their most inter-connected protein-coding genes. Information from the disease module genes may give insights into the underlying disease mechanisms. We used high confidence interactions from STRING-db (version 10.5, confidence score>0.7)⁹ as source of the PPI network, as it includes several different data-bases and was found to be the best network database for disease gene discovery in a systematic evaluation of 21 different molecular networks.¹⁰

Protein IDs were converted to corresponding gene symbols resulting in a network with gene-gene interactions. In order to be less biased by module identification method we employed four module defining methods, namely: Clique-sum¹¹, DIAMOnD¹², MCODE¹³ and ModuleDiscoverer¹⁴, each identifying disease modules using our DMP-genes as seed genes. The four methods were utilised through the R package MODifieR (version 0.1.0, available at: <https://gitlab.com/Gustafsson-lab/MODifieR>). Each method yielded a list of genes corresponding to an identified module. From these four methods, a consensus module was obtained by extracting those genes found in modules from at least two of the methods. Network visualisation was conducted in Cytoscape (version 3.5.1).¹⁵

Codice campo modificato

398

399

400

401 **GO-enrichment in relation to development of atopic manifestations and the**
402 **construction of a classifier**

403 Next, the biological relevance of the module genes was assessed through pathway
404 enrichment analysis. Based on biological knowledge from Gene Ontology (GO), biological
405 themes significantly overrepresented among the genes were identified using the R package
406 clusterProfiler (version 3.4.4).²¹ Each of the identified clusters in the module network were
407 tested for both GO- and KEGG-enrichment using the online tools of the Database for
408 Annotation, Visualization and Integrated Discovery (DAVID, version 6.8).^{19,20} In addition, the
409 module genes were linked to allergy outcome in a linear model using the limma package for
410 R (version 3.32.10). This was done by first extracting probes annotated for all module genes,
411 that is for genes identified from the initial DNA methylation analyses, as well as for genes
412 deriving from the PPI-network, rendering 2743 probes that were studied further. Thereafter, a
413 linear model was fitted to the respective methylation values at birth, 12 and 24 months, using
414 sampling time points as co-variables. Comparisons were made between individuals
415 developing atopic manifestations (sensitisation and/or allergic symptoms, cumulatively up to
416 7 years of age) and individuals who did not. Furthermore, the power of the module (n=91,
417 probiotics vs. placebo at birth), to predict which children would go on to develop atopic
418 manifestations (n=24, nProb=12, nPlac=12) and which were going to remain healthy (n=17,
419 nProb=5, nPlac=12) at seven years of age, was investigated by testing the extracted module
420 methylation probes (n=2734). Gene enrichment on genes associated with allergies was
421 performed, followed by the construction of a classifier. Various standard robust classification
422 methods were tested from the Matlab classification toolbox (with default parameters), with
423 the best method being Ensemble learning (KNN). All predictions were derived using a leave-
424 one-out cross-validation, and to account for stochastic fluctuations the learning inference
425 procedure was repeated 100 times.

426

427

428

429

430

431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447

PCR programme specifications

Primer sequences used in locus-specific PCR and SNaPshot reaction

Sequences – locus-specific primers:	
IL13	Forward: TTGGGTGATGTTGATTAGTTTTTTAATGAG Reverse: CAAATCTTAAAACTCTACCCTAAACCC
FOXP3	Forward: TGGTGAAGTGGATTGATAGAAAAGG Reverse: TATAAAAACCCCCCCCACC
IFNG	Forward: TGGTTAGAAGGTATAAAAGAAAAGG Reverse: ATCACATCCTAACACTCACAAC
IL17	Forward: GAGATTTTTTTATGATTTTATTG Reverse: TTATACCTACTATAAAATAAACAA
Sequences - SNaPshot primers:	
IL13	5'-TTTTTTTTTTGTAGTTTTTATAGGTTTAAGTGGTGA-3'
FOXP3	5'-AAAAAAAAAAAAATTTAAATTATTAAAAAAAAAAAAATCTAC-3'
IFNG	5'- GATTTATATTGAGTGGAAGAAATTTA-3'
IL17	5'- CACCATTTTAATTAATTTCC-3'

Chromosome location and position of investigated loci in relation to transcription start site (TSS): *IL13* (5q31.1, position -46 TSS)⁵, *IL17* (6p12.2, position -122 TSS)⁶, *FOXP3* (Xp11.23, position -77 TSS)⁷ and *IFNG* (12q15, position -4229 TSS)⁴.

448

449

450

451 **PCR programmes for T helper subset locus-specific DNA methylation analysis****Bisulphite conversion**

98°C	10 min		
64°C	3.5 h		
4°C	storage		

Locus-specific PCR

96°C	2 min		
95°C	30 sec Annealing temperature*, 30 sec	35 cycles	
72°C	1 min		
72°C	5 min		
4°C	storage		

*Annealing temperatures: 49°C for *IL-17* locus and 53.1°C for *IL-13*, *FoxP3* and *IFN-γ* loci.

ExoStar			
37°C	15 min		
80°C	1 h		

Snapshot

96 °C	10 sec		
50 °C	5sec30		
	cycles		
60 °C	30 sec		
4 °C	storage		

CIP			
-----	--	--	--

37°C	1h		
75°C	15 min		

References:

1. Abrahamsson TR, Jakobsson T, Böttcher MF, et al. Probiotics in prevention of IgE-associated eczema: a double-blind, randomized, placebo-controlled trial. *J Allergy Clin Immunol*. 2007;119(5):1174-1180.
2. Abrahamsson TR, Jakobsson T, Björkstén B, Oldaeus G, Jenmalm MC. No effect of probiotics on respiratory allergies: a seven-year follow-up of a randomized controlled trial in infancy. *Pediatr Allergy Immunol*. 2013;24(6):556-561.
3. Forsberg A, Abrahamsson TR, Björkstén B, Jenmalm MC. Pre- and postnatal administration of *Lactobacillus reuteri* decreases TLR2 responses in infants. *Clin Transl Allergy*. 2014;4:21.
4. Janson PC, Marits P, Thorn M, Ohlsson R, Winqvist O. CpG methylation of the IFNG gene as a mechanism to induce immunosuppression [correction of immunosuppression] in tumor-infiltrating lymphocytes. *J Immunol*. 181(4).
5. Santangelo S, Cousins DJ, Winkelmann NEE, Staynov DZ. DNA methylation changes at human Th2 cytokine genes coincide with DNase I hypersensitive site formation during CD4(+) T cell differentiation. *Journal of immunology (Baltimore, Md : 1950)*. 2002;169(4):1893-1903.
6. Janson PC, Linton LB, Bergman EA, et al. Profiling of CD4+ T cells with epigenetic immune lineage analysis. *J Immunol*. 186(1).
7. Janson PC, Winerdal ME, Marits P, Thorn M, Ohlsson R, Winqvist O. FOXP3 promoter demethylation reveals the committed Treg population in humans. *PLoS one*. 2008;3(2).
8. Sobrino B, Brión M, Carracedo A. SNPs in forensic genetics: a review on SNP typing methodologies. *Forensic science international*. 2005;154(2-3):181-194.
9. Benjamini Y, Krieger AM, Yekutieli D. Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*. 2006;93(3).
10. Morris TJ, Butcher LM, Feber A, et al. ChAMP: 450k Chip Analysis Methylation Pipeline. *Bioinformatics (Oxford, England)*. 2014;30(3):428-430.
11. Teschendorff AE, Marabita F, Lechner M, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics (Oxford, England)*. 2013;29(2):189-196.
12. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC bioinformatics*. 2012;13:86.
13. Morin AM, Gatev E, McEwen LM, et al. Maternal blood contamination of collected cord blood can be identified using DNA methylation at three CpGs. *Clinical Epigenetics*. 2017;9(1):75.
14. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics (Oxford, England)*. 2007;8(1):118-127.
15. Weiner J. PCA 3D: getting PCA plots quickly. *PCA 3D: getting PCA plots quickly*. 0100.
16. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research*. 2015;43(7).
17. Van der Most PJ, Küpers LK, Snieder H, Nolte I. QCEWAS: automated quality control of results of epigenome-wide association studies. *Bioinformatics (Oxford, England)*. 2017;33(8):1243-1245.
18. van Iterson M, van Zwet EW, Consortium B, Heijmans BT. Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution. *Genome Biology*. 2017;18(1):19.
19. Huang DWaW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*. 2009;4(1):44-57.
20. Huang DWaW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic acids research*. 2009;37(1):1-13.
21. Yu G, Wang L-GG, Han Y, He Q-YY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics : a journal of integrative biology*. 2012;16(5):284-287.

