

Contrasting microbiotas between Finnish and Estonian infants: Exposure to *Acinetobacter* may contribute to the allergy gap

Lasse Ruokolainen¹  | Anna Parkkola^{2,3} | Antti Karkman^{1,4} | Hanna Sinkko^{5,6} | Aleksandr Peet⁷ | Anu-Maaria Hämäläinen⁸ | Leena von Hertzen⁹ | Vallo Tillmann⁷ | Katriina Koski¹⁰ | Suvi M. Virtanen^{11,12,13} | Onni Niemelä¹⁴ | Tari Haahtela⁹  | Mikael Knip^{2,3,13,15} 

¹Department of Biosciences, University of Helsinki, Helsinki, Finland

²Pediatric Research Center, Children's Hospital, Helsinki University Hospital, University of Helsinki, Helsinki, Finland

³Research Program for Clinical and Molecular Metabolism, Faculty of Medicine, University of Helsinki, Helsinki, Finland

⁴Department of Microbiology, University of Helsinki, Helsinki, Finland

⁵Human Microbiome (HUMI) Research Program, Faculty of Medicine, University of Helsinki, Helsinki, Finland

⁶Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

⁷Department of Paediatrics, Tartu University Hospital, University of Tartu, Tartu, Estonia

⁸Department of Paediatrics, Jorvi Hospital, Helsinki University Hospital, Espoo, Finland

⁹Skin and Allergy Hospital, Helsinki University Hospital, Helsinki, Finland

¹⁰Clinicum, Faculty of Medicine, University of Helsinki, Helsinki, Finland

¹¹Department of Public Health Solutions, National Institute for Health and Welfare, Helsinki, Finland

¹²Faculty of Social Sciences/Health Sciences, University of Tampere, Tampere, Finland

¹³Center for Child Health Research, Tampere University Hospital and Science Centre, Tampere University Hospital, University of Tampere, Tampere, Finland

¹⁴Department of Laboratory Medicine and Medical Research Unit, Seinäjoki Central Hospital, University of Tampere, Tampere, Finland

¹⁵Folkhälsan Research Center, Helsinki, Finland

Correspondence

Mikael Knip, Children's Hospital, University of Helsinki, P.O. Box 22, Helsinki 00014, Finland.

Email: mikael.knip@helsinki.fi

Funding information

Medicinska Understödsföreningen Liv och Hälsa; Terveyden Tutkimuksen Toimikunta, Grant/Award Number: 250114; FP7 Health, Grant/Award Number: 202063

Abstract

Background: Allergic diseases are more common in Finland than in Estonia, which—according to the biodiversity hypothesis—could relate to differences in early microbial exposures.

Methods: We aimed at defining possible microbial perturbations preceding early atopic sensitization. Stool, nasal and skin samples of 6-month-old DIABIMMUNE study participants with HLA susceptibility to type 1 diabetes were collected. We compared microbiotas of sensitized (determined by specific IgE results at 18 months of age) and unsensitized Estonian and Finnish children.

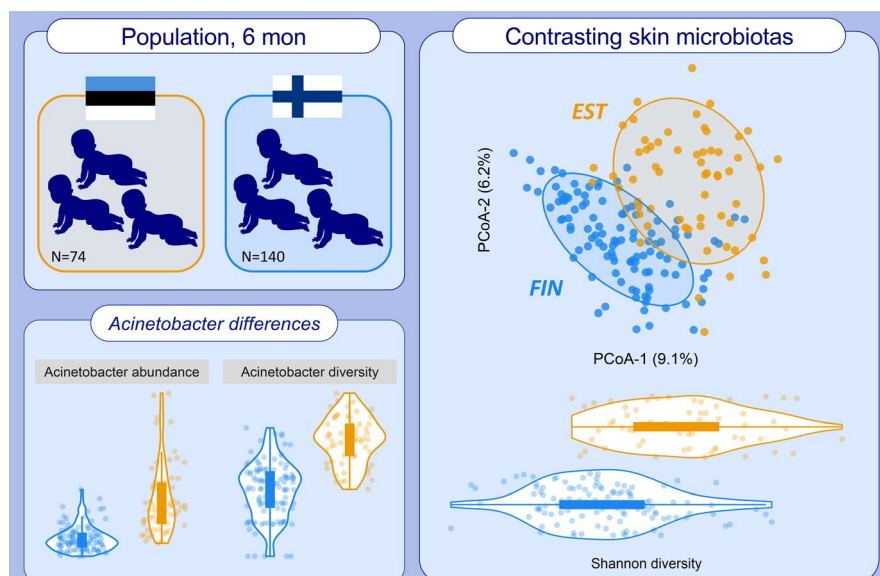
Results: Sensitization was differentially targeted between populations, as egg-specific and birch pollen-specific IgE was more common in Finland. Microbial diversity and community composition also differed; the genus *Acinetobacter* was more abundant in Estonian skin and nasal samples. Particularly, the strain-level profile of *Acinetobacter lwoffii* was more diverse in Estonian samples. Early microbiota was not generally associated with later sensitization. Microbial composition tended to differ

between children with or without IgE-related sensitization, but only in Finland. While land-use pattern (ie green areas vs. urban landscapes around the children's homes) was not associated with microbiota as a whole, it associated with the composition of the genus *Acinetobacter*. Breastfeeding affected gut microbial composition and seemed to protect from sensitization.

Conclusions: In accordance with the biodiversity hypothesis, our results support disparate early exposure to environmental microbes between Finnish and Estonian children and suggest a significant role of the genus *Acinetobacter* in the allergy gap between the two populations. The significance of the observed differences for later allergic sensitization remains open.

KEYWORDS

atopic sensitization, intestine, microbiota, respiratory, skin



GRAPHICAL ABSTRACT

Estonian children had a more diverse intestinal microbiota than Finnish peers at the age of 6 months. Microbial community composition differed significantly between 6-month-old Estonian and Finnish children. Especially, genus *Acinetobacter* was more abundant in skin- and nasal samples of Estonian children. Our results support the importance of early exposure to environmental microbes and suggest that exposure to *Acinetobacter* might be particularly protective.

1 | INTRODUCTION

Immune-mediated diseases such as allergies and autoimmune diseases represent one possible cost of our urbanized lifestyle. Reduced contact with diverse microbiotas of natural environments contributes to the rising epidemic of immune-mediated diseases, as explained by the biodiversity hypothesis of allergic disease.¹ Many factors, such as living on a farm, having older siblings and having pets, have been linked to decreased risk for allergy and atopy.²⁻⁴ All of these factors, which either vanish or become less likely with an urban lifestyle, could be traced back to increased exposure to environmental microbes,⁵ which potentially translates into increased

microbial colonization to the body. Importantly, some of these environmental microbes have been linked to increasing immune tolerance.⁵⁻⁷

A living laboratory to test the biodiversity hypothesis exists across the Finnish borders. Westernized living conditions in Finland meet a more traditional lifestyle on the Russian side. Studying these contrasting populations, a three-fold to ten-fold increased prevalence of asthma or allergic sensitization,^{4,8} a four-fold increased prevalence of coeliac disease,⁹ and a six-fold increased incidence of type 1 diabetes has been reported in Finland,¹⁰ as compared to Russian Karelia. Estonia is a country of intermediate level of westernization compared to Finland and Russia. It has also witnessed a rapid socioeconomic growth since its independence in 1991. Type

1 diabetes and coeliac disease are three times more common in Finland compared to Estonia,^{11,12} and wheezing during the past year has been reported in 7%-10% of Estonian compared to 15%-20% of Finnish children.^{13,14} Similarly, atopic dermatitis and skin prick test positivity are more prevalent among Swedish infants (24%) as compared to Estonian infants (~14%).¹⁵ According to the DIABIMMUNE study, allergic sensitization to any allergen by the age of 3 years was more common in Finland compared to Estonia (45.6% vs. 32.4%, $P = .001$; Mustonen N, personal communication).

These reported differences in epidemiology have been shown to parallel differences in the human microbiotas,^{8,16} which is likely to reflect differential exposure to environmental microbes. Gut microbial composition has been shown to differ significantly between Finnish and Russian children, but not substantially so between Finnish and Estonian children.¹⁶ Also, skin and nasal microbiotas were contrastingly different in composition between adolescents from Russian and Finnish Karelia.⁸ This is in line with many other observations that dysbiosis—abnormal/imbalanced microbiota—is associated with the development of inflammatory disorders.¹⁷

Many previous studies have compared either school children or young adults. However, environmental exposures crucial to later development of allergic diseases may occur during early life.^{3,18} Accordingly, in this study we compared skin, nasal mucosa and gut microbiotas from infants from Finland and Estonia. Our aim was to search for early differences in the commensal microbiota that might predict early allergic sensitization and the gap in the prevalence of allergic diseases between these two populations.

2 | MATERIALS AND METHODS

2.1 | Subjects and sample collection

The DIABIMMUNE birth cohort (BC) was a prospective study in Finland, Estonia and Russia following children with HLA susceptibility to type 1 diabetes from birth to the age of 3 years from September 2008 to October 2013.¹² Blood and swab samples were collected at the study centre visits (3, 6, 12, 18, 24 and 36 months of age).

For this study, children observed to be sensitized according to the IgE results by the age of 18 months were included from the Finnish ($n = 387$) and Estonian DIABIMMUNE cohorts ($n = 330$). Unfortunately, not enough Russian samples were available for this study. The allergen-specific IgE levels tested at 18 months of age were cat, hen's egg, cow's milk, peanut, birch, timothy and dust mite (*Dermatophagoides pteronyssinus*). To include only children with a definite allergic sensitization, a cut-off of IgE ≥ 0.7 kU/L in at least one serum allergen-specific IgE level was used for the identification of cases instead of the standard cut-off of 0.35 kU/L. In the statistical analyses, we classified subjects as sensitized, based on an empirical cut-off of the maximum specific IgE value, set at -0.5 on $\log_{10}$ scale (Figure S1). Table S1 shows the proportion of Finnish and Estonian study participants with positive sIgE responses (≥ 0.35 kU/L) to the seven allergens tested at the age of 18 months.

The inclusion criteria resulted in the identification of 70 Finnish children (18.1%) but only 28 Estonian children (8.5%) with allergic sensitization, reflecting the existing gap in allergy prevalence between these two populations. For the 70 Finnish children, we included 70 country-matched controls with no allergen-specific IgE levels above 0.35 kU/L at the age of 18 months. Due to the lower number of sensitized children in Estonia, we included two controls for the 28 Estonian children. In total, this study comprised 224 children (117 boys), 98 children with definite allergic sensitization by allergen-specific IgE levels at 18 months of age and 126 controls. Table S2 presents the demographic and other characteristics of the Finnish and Estonian study participants, and Table S3 presents the corresponding characteristics in the sensitized and nonsensitized participants.

Stool, skin and nasal samples were collected at the age of 6 months. Nasal swab samples were available for 217 children ($n_{\text{Fin}} = 135$, $n_{\text{Est}} = 82$), skin swab samples for 172 children ($n_{\text{Fin}} = 111$, $n_{\text{Est}} = 61$) and stool samples for 218 children ($n_{\text{Fin}} = 135$, $n_{\text{Est}} = 83$). A complete set of samples was obtained from 162 children ($n_{\text{Fin}} = 102$, $n_{\text{Est}} = 60$). For those children without a stool sample at 6 months of age, a sample at 5 months was selected ($n = 14$), and for those lacking both samples, a sample at 7 months was selected ($n = 2$).

The local ethics committees approved the DIABIMMUNE study protocol, and legal guardians of the study participants gave their written informed consent.

2.2 | Land-use patterns

Land use was calculated with the CORINE2012 land-cover data, using five land-use types: agricultural land, built area, forest, water bodies and wetland. We used a buffer with a radius of 3 km around the homes.¹⁹

2.3 | DNA extraction and sequencing

Bacterial composition was determined after DNA extraction by amplifying V1-V3 regions of bacterial 16S rRNA genes by polymerase chain reaction (PCR) (see Appendix S1).

2.4 | Bioinformatics

The bioinformatics methods have been described in the Appendix S1.

2.5 | Statistics

We tested the ability of microbial composition to classify individuals according to national identity, using random forest analysis, as implemented in the randomForest package.²⁰ This analysis was also

used to identify characteristic operational taxonomic units (OTUs) of each group. Prior to analysis, in order to limit the number of feature included in the analysis, we selected only those OTUs having a variance higher than the median variance across all OTUs. An optimal number of trees were selected based on classification error, and an optimal number of randomly selected variables used in the analysis were sought using the train function in the caret package,²¹ based on 10-fold cross-validation. Due to imbalanced sample sizes—larger sample size of Finnish children would skew the prediction so that the Estonian children could not be predicted reliably, we took a random sample of $n_{\text{skin}} = 60$, $n_{\text{nasal}} = 80$ and $n_{\text{stool}} = 80$ samples from each population to be used in the analysis. The number of features used was 1391 for skin and nasal microbiotas, and 1387 for stool microbiota. For skin, nasal and stool samples, we used 100, 100 and 200 trees, respectively, and 100, 197 and 100 random features, respectively.

Questionnaire data on animal contact were converted to binary form (by coding high/prominent contact with 1 and low/infrequent contact with 0). The resulting data were used to calculate an aggregate estimate of the intensity of animal contact, with principal component analysis (PCA), where the first PCA axis was used in the analysis.

A more detailed description of the methods can be read in the Appendix S1.

3 | RESULTS

3.1 | Between-population differences in allergic sensitization

For the statistical analyses, we used the maximum specific IgE values, measured at the age of 18 months, to classify individuals to either “healthy” or “sensitized,” setting the cut-off at -0.5 on $\log_{10}$ scale (0.61 on arithmetic scale; based on the frequency distribution of IgE values, Figure S1). The maximum IgE at 18 months correlated relatively well with that at 6 months ($n_6 = 210$; Spearman $\rho = 0.69$, $P < .0001$) and 36 months ($n_{36} = 198$; Spearman $\rho = 0.74$, $P < .0001$). These correlations did not differ markedly between population cohorts.

The IgE values of individual allergens were generally very low; the medians of all measured allergens were less than 0.1 kU/L. As low measures below 0.1 kU/L are not considered to be trustworthy, we used this level as cut-off. The data suggest that egg-specific and birch-specific IgE is elevated in Finland (egg: $OR_{\text{Finland}} = 2.22$, $CI = [1.25, 3.95]$, $P = .0068$; birch: $OR_{\text{Finland}} = 4.69$, $CI = [1.04, 21.17]$, $P = .045$). Interestingly, house dust mite-specific IgE was elevated in Estonia, but the difference was not statistically significant ($OR_{\text{Finland}} = 0.36$, $CI = [0.11, 1.14]$, $P = .083$). These results suggest that the immune response might be differently targeted between the populations, even at this early age, which is in accordance with previous observations between Finnish and Russian Karelia.⁸

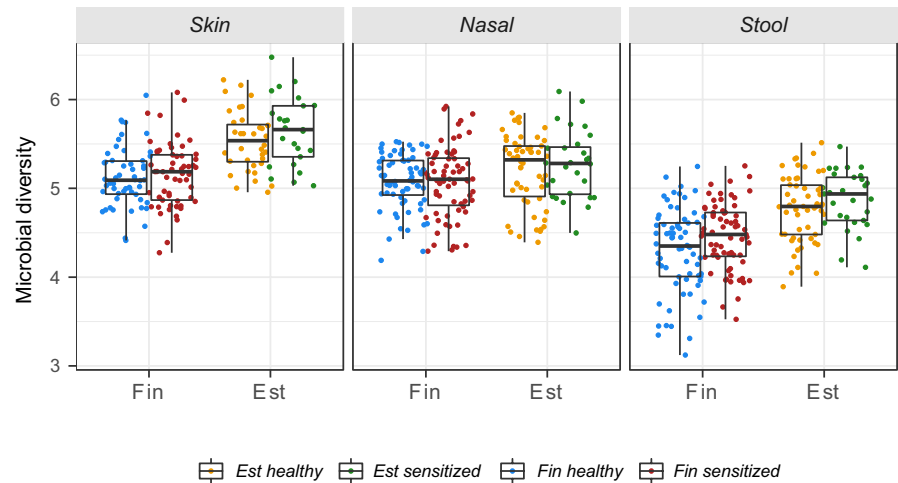
3.2 | Patterns in the composition of microbiota

The sequencing resulted in 2782 OTUs shared between both sequencing runs in the 607 samples from skin, nasal and stool, obtained at the age of 6 months. Microbial diversity was highest in the skin samples ($P = .0004$) and lowest in the stool samples ($P < .0001$) (Figure 1). Across all samples, Estonian children tended to have higher microbial diversity, as compared to Finnish children ($P < .0001$), and especially so when considering either skin or stool samples. Considering diversity between the populations, as well as healthy and sensitized children (assessed at 18 months), the only significant difference was that the Estonian children had more diverse microbiota in their stool compared to Finnish children ($P = .007$). In the skin samples from Estonia, diversity tended to be higher in sensitized children as compared to healthy children ($P = .028$) (Figure 1).

Microbial community composition differed most clearly between populations when considering skin samples (multiple regression on distance matrices (MRM) on weighted UniFrac: $R^2 = 0.08$, $P < .001$; Figure 2A), followed by the nasal samples ($R^2 = 0.03$, $P < .001$; Figure 2B), and stool samples ($R^2 = 0.02$, $P < .001$; Figure 2C). Significant differences between healthy and sensitized children were observed both in the nasal and stool samples, but only among Finnish children (MRM on Bray-Curtis: nasal: $R^2 = 0.02$, $P < .001$; stool: $R^2 = 0.01$, $P = .013$). These analyses were corroborated by random forest classification (RFC). The overall classification accuracy (based on 10-fold cross-validation of a random subsample of the data) was highest for skin microbiota (accuracy = 92.5%), followed by nasal (accuracy = 91.9%) and stool microbiota (accuracy = 81.3%). In all cases, subject recall was higher for Finnish than for Estonian subjects: nasal samples Fin = 92.5%, Est = 91.3%; skin samples Fin = 95%, Est = 90%; and stool samples Fin = 86.3%, Est = 76.3%.

Using RFC, we were able to assess the most important OTUs differentiating between Finnish and Estonian children (Figure 2.). In both nasal and skin samples, the most important OTU belonged to genus *Acinetobacter* (OTU-108) (Figure 2A, B). This OTU was more abundant in Estonian children as compared to Finnish children—on average over 20 times more abundant on the skin ($P_{\text{skin}} < .0001$) and four times more abundant on the nasal epithelium ($P_{\text{nasal}} = .0025$) (Figure 3A). Furthermore, it could be identified to species *A. lwoffii* (a blast search against the NCBI 16S ribosomal RNA sequence database gave a 99% match over 486 nucleotides). The “strain”-level profile was also found to be different between Finnish and Estonian children using an oligotyping pipeline ($R^2 = 0.06$, $P < .001$, based on Bray-Curtis dissimilarity, correcting for sample type; Figure 3B). Again, the strain profile of OTU-108 was more diverse (measured by Shannon diversity) among Estonian children (P_{nasal} and $P_{\text{skin}} < .0001$). While the strain profile seems to differ especially between healthy and sensitized Finnish children, these differences were not statistically significant.

FIGURE 1 Microbial diversity (Shannon's index) in different sample types (skin, nasal and stool), comparing healthy and sensitized subjects among the Finnish and Estonian cohorts



3.3 | Living environment affecting skin and nasal microbiotas

Land-use patterns, characterized as the first principal component of land-use types, differed significantly between populations ($P = 9.1e-8$); Estonian subjects lived in more rural surroundings than the Finnish subjects. This is reflected in the analysis of dissimilarities, such that for both skin and nasal microbiotas, land use did not remain significant after correcting for the significant effect of population ($P = .001$). As reported before,¹⁹ land use was not associated with sensitization in these data. Also, (taking a random sample of subjects from each population: $n_{\text{skin}} = 50$, $n_{\text{nasal}} = 70$), only 18% of variation along the land-use gradient could be predicted with skin microbiota (using 150 trees and 225 features) and 11% based on nasal microbiota (using 150 trees and 500 features), using random forest regression. Interestingly, while land use was not associated with microbiota as a whole, there was a significant association between *Acinetobacter* composition in the nasal samples and land use. After accounting for the population effect ($P = .001$), both the amount of agricultural land ($P = .002$) and forest cover ($P = .029$) were associated with similarity in *Acinetobacter* composition between samples (Bray-Curtis dissimilarity on sqrt transformed OTU-level counts).

3.4 | Stool microbiota and other lifestyle factors

To account for possible dietary effects, we considered the influence of breastfeeding patterns on stool microbiota. While the duration of breastfeeding did not differ between populations (Table S3), those children who were exclusively breastfed for at least the first month of life were less likely to be sensitized at the age of 18 month ($P = .00015$), after accounting for population, mode of delivery and the month to start on solid foods (Table 1). This could not be explained by between-population differences in exclusive breastfeeding ($P = .59$), indicating that this protective effect occurs across populations. Interestingly, stool microbial diversity was significantly

reduced by breastfeeding at the time of sampling ($P < .0001$), again accounting for population, mode of delivery and the month to start on solid foods (Table 1). Moreover, breastfeeding at time of sampling was also significantly associated with variation in between-sample (weighted) UniFrac distances ($P = .002$; Table 1), indicating that breastfeeding influenced the composition of the gut microbial community. Particularly, breastfeeding was associated with an increase in the abundance of Lactobacillaceae ($P = .0003$, when accounting for population effect).

Finally, we tested whether the amount of animal contact in early life (6 months) was associated with later sensitization and microbial composition (measured at 18 months). While animal contact was significantly more intensive in Estonia as compared to Finland ($P = 1.44e-7$), it was not associated with allergic sensitization ($P = .94$), after accounting for the between-population difference ($P = .041$). However, stool microbiota was significantly associated with animal contact (MRM on weighted UniFrac, $R^2 = 0.02$, $P = .003$; corrected for between-population differences), which might be reflected in later health. When comparing individuals with low (1st quartile) and high (4th quartile) animal contact, a few OTUs identified by RFC were significantly associated with animal contact. These OTUs represented the genera of *Anaerostipes* (Kruskal-Wallis test, $P = .0023$), *Blautia* ($P = .0074$) and *Bacteroidetes* ($P = .014$), all more abundant in the high-contact group.

4 | DISCUSSION

In this study of 6-month-old Finnish and Estonian infants, our objective was to compare patterns in microbiota between populations and discover potential microbial predictors of early sensitization to allergens. Indeed, Finnish and Estonian infants were found to differ in their skin, nasal and stool microbiota. The role of these differences for later developing allergic sensitization could not be confirmed at this stage. However, the results clearly indicate—in agreement with earlier findings—that observed between-population differences in epidemiology could be partly explained by differential exposure to environmental microbiota. Such a conclusion is supported by the

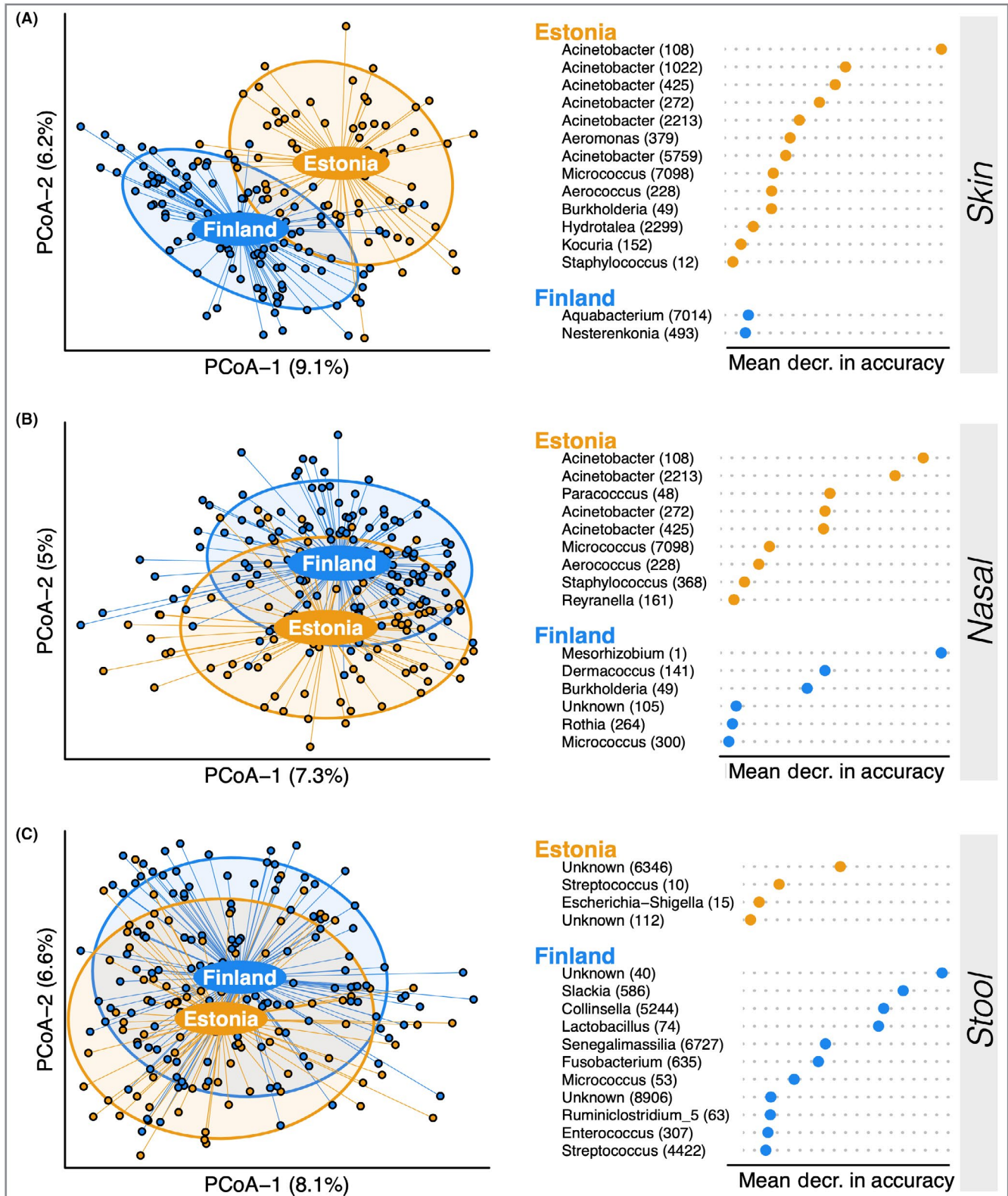
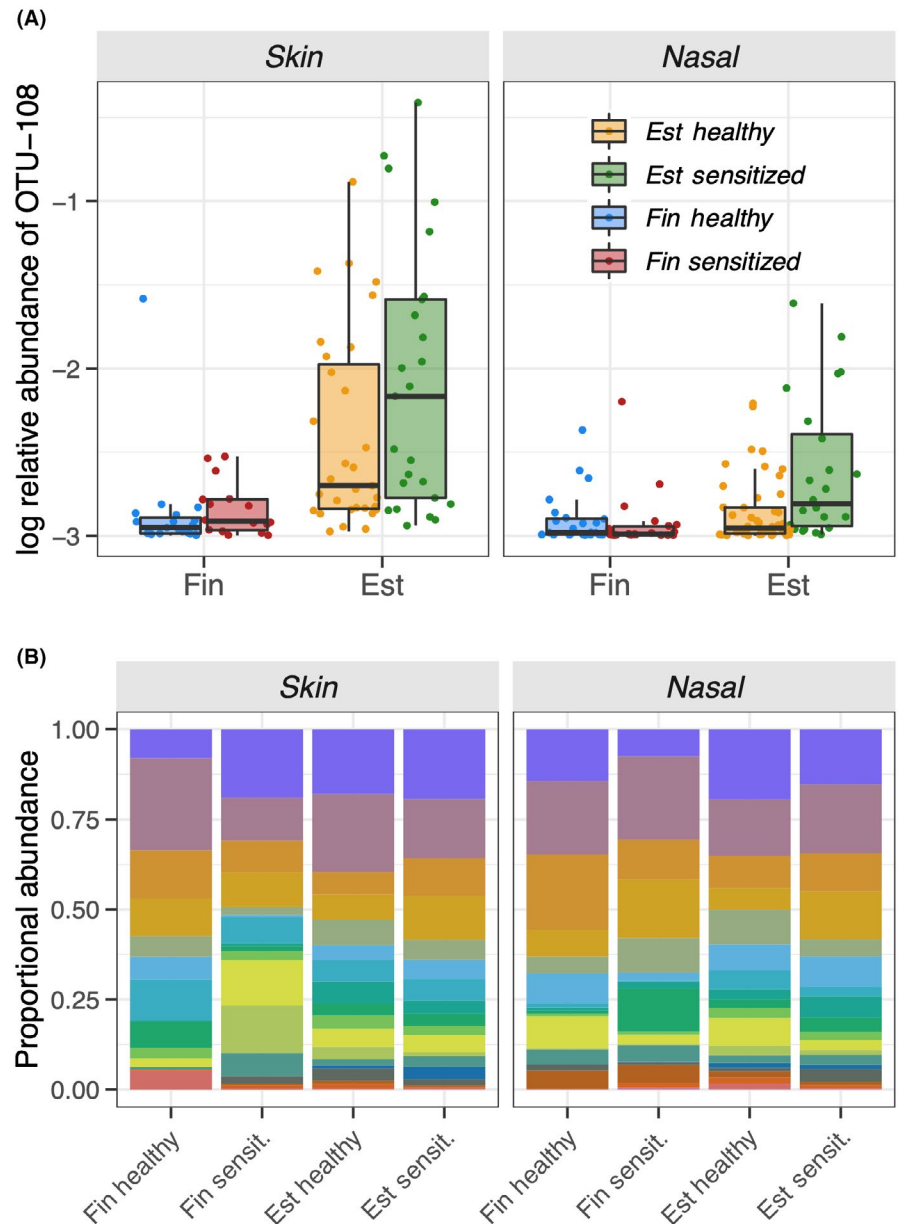


FIGURE 2 Ordination of community composition and the top discriminative OTUs for each cohort in either the (A) skin, (B) nasal or (C) stool samples. Most important OTUs differing between Finnish and Estonian children were found with random forest classification. Ordinations were based on the Bray-Curtis dissimilarity

fact that the Estonian cohort is exposed to more rural environment, and they also tend to be in closer contact with farm animals as compared to the Finnish cohort.

The contrasting microbial community composition between Finnish and Estonian children on the skin and the nasal epithelium was best characterized by differential abundance of several OTUs from

FIGURE 3 (A) Proportional abundance (B) and the oligotype profile of OTU-108 (*Acinetobacter lwoffii*) in healthy and sensitized Finnish and Estonian children, in both skin and nasal samples. An oligotype is a group of reads that are binned together based on the nucleotides they possess by using Shannon entropy



the genus *Acinetobacter* (Figure 2A, B). In particular, we found that *A lwoffii* (and the genus in general) was substantially more abundant, and genetically more diverse, in the skin and nasal samples from Estonian than Finnish infants (Figure 3). This resonates with our earlier observation that adolescents in the Russian Karelia have significantly higher abundance of more diverse *A lwoffii* on their skin and nasal epithelium than their Finnish counterparts.⁸ Importantly, the present results indicate that these differences are detectable already during infancy and precede the development of allergic sensitization. This suggests sustained exposure to *Acinetobacter* in these populations, which is likely to begin already before birth and range into adulthood.

Within the genus *Acinetobacter*, *A lwoffii* seems to be the most common commensal on human skin and nasal mucosa.^{22,23} These bacteria are found in soil,²⁴ but they are also relatively abundant in cowshed microflora.⁷ This indicates that the living environment and lifestyle might contribute to the colonization of this bacterium to the

body, which could explain the observed between-population differences. This is in turn supported by the previous observation that exposure to *A lwoffii* in the living environment is associated with reduced risk of allergic sensitization²⁵. In the present context, it would be extremely interesting to follow health outcomes of the participants in later life to see whether this is indeed the case.

Remarkably, *A lwoffii* has been shown to polarize the immune system towards Th1 responses in a mouse model.^{7,26} Furthermore, even prenatal exposure to *A lwoffii* can significantly prevent the development of an asthmatic phenotype in the progeny in the mouse model.^{27,28} For example, Brand et al²⁸ showed that the exposure of the mother to *A lwoffii* leads to epigenetic modifications in the progeny, promoting the production of IFN- γ . More recently, Fyhrquist et al²⁶ demonstrated that exposure to *A lwoffii* through the skin is associated with elevated expression of IL-10 and IFN- γ , and suppression of IL-5 and IL-13 expression. This local Th1-polarization of

TABLE 1 (A) Exclusive breastfeeding at 1 month of age reduced the risk (OR) of allergic sensitization (logistic regression model). (B) Breastfeeding at the time of sampling reduced microbial (Shannon) diversity (generalized least squares model, account for between-group residual variation) and community composition (weighted UniFrac distance, permutational regression on distance matrices) in stool samples

Response	Predictor	OR/Estimate/R ²	P
(A) Sensitization	Country (Estonia)	0.47	.018
	Delivery mode (CS)	0.47	.12
	Number of siblings	0.67	.009
	Month to start solid food	1.13	.45
	Exclusive breastfeeding at 1 mon	0.29	.0003
(B) Diversity	Country (Estonia)	0.40	<.0001
	Delivery mode (CS)	-0.10	.70
	Number of siblings	-0.01	.30
	Month to start solid food	-0.06	.0001
	Breastfeeding at 6 mon	-0.30	<.0001
(C) Composition	Country	0.02	.003
	Delivery mode	0.02	.002
	Number of siblings	0.01	.14
	Month to start solid food	0.02	.003
	Breastfeeding at 6 mon	0.024	.001

Note: All analyses were controlled for country, delivery mode (vaginal or caesarean section = CS), number of siblings and the month when solid foods were included in the diet.

the immune response was also associated with systemic protection from allergic sensitization and inflammation on allergen challenge. These experimental findings, together with our observations, suggest that the Estonian and Russian children are being protected against allergic sensitization via both pre- and postnatal exposure to *Acinetobacter*.

Also, the faecal samples at the age of 6 months differed between populations, but the differences were less clear. While the relatively small between-population difference is understandable due to the selectivity of the gut habitat,^{17,29} Estonian children had a more diverse gut microbiota compared to Finnish children. Lower diversity of gut microbiota has been associated with the development of allergic diseases,³⁰ which is in line with our findings of reduced diversity in a population sampled from a country with

higher prevalence of allergic diseases. However, it is not clear why this difference arises. In both cohorts, breastfeeding was found to be protective against sensitization. Breastfeeding was also associated with an increase in the abundance of Lactobacillaceae, in agreement with previous reports.³¹ This could relate to the protective role of breastfeeding via increased colonic production of short-chained fatty acids,³² which can shape the immunological environment in the lung and influence the severity of allergic inflammation.³³ Notably, the compositional change in gut microbiota in the context of breastfeeding was also associated with reduced microbial diversity. We interpret this as the functional role of the microbial community members being more important for the host than mere diversity.

We searched for predictors of early allergic sensitization at the age of 18 months and found the detected differences in microbiotas to be somewhat surprising. Microbial diversity tended to be higher in sensitized children as compared to healthy children (Figure 1), although these differences were not statistically significant. Additionally, *A. lwoffii* tended to be more abundant in sensitized compared to healthy infants (Figure 3). Based on the assumption that higher microbial diversity is associated with protection from allergic diseases and the fact that *A. lwoffii* being more abundant in Estonian samples was proposed to play a role in a lower allergy prevalence in this population, one might have expected to see an opposite pattern. It should be kept in mind that it is not clear that higher microbial diversity in the gut per se would be beneficial.¹⁷ Patterns in diversity likely arise from interactions between taxa, which are affected by factors such as diet and medication.¹⁷

Another possible explanation could be that our selected endpoint, sensitization at 18 months of age, was not detecting true atopic development in our study population but rather reflected exposure of these infants to different allergens. Consequently, a child displaying an IgE response to allergens could merely have been exposed to a larger array or higher loads of these allergens, which might also be evident as a more diverse microbiota. According to Lynch et al,³⁴ such a scenario can actually protect against, rather than predispose to, allergic disease. It is possible that at the age of 18 months, the children sensitized do not necessarily represent children who will later develop allergic diseases but rather this IgE response will later dissolve. For example, in early life, sensitization towards food allergens is more frequent (here, the maximum IgE at 18 months came from a food allergen in 78% of the subjects) than in later life, and for example in a German study, two-thirds of the children with sensitization by the first year of life did not develop any atopic disease by the age of six years.³⁵

Also, sensitization to dust mite—especially monosensitization—might not be related to later atopic diseases but could even provide protection from such conditions.³⁶ Only three children, however, one Estonian and two Finnish, were sensitized to dust mite at the level of 0.35 kU/L. None of these children were monosensitized, although dust mite IgE levels higher than 0.1 kU/L were more common

among Estonian than Finnish children. However, breastfeeding was a significant factor potentially contributing to the development of sensitization; infants exclusively breastfed for at least one month were less likely to be sensitized at the age of 18 months. While this corroborates many earlier studies describing breastfeeding as a protective factor for allergic diseases,³⁷ there are also several studies that do not agree.^{38,39}

A limitation of our study is that the included children represent selected cases in terms of HLA class II genetics. Accordingly, the results might not directly be generalizable to the whole population, and possibly, some differences existing in the whole population are missed. We compared microbial samples at the age of 6 months, which in the light of recent findings might already be late. Especially, the alterations in stool microbiota have often been described very early in infancy: Arrieta et al¹⁸ reported a transient intestinal microbial dysbiosis at 3 months of age in infants at high risk of later asthma development, which had mostly disappeared already at 1 year of age. Even earlier differences in association with later development of allergic disease have been reported,^{30,40} and in the DIABIMMUNE cohort, also differences in umbilical cord blood transcriptomics have been described.⁴¹

Another limitation is that the number of subjects is small. This practically prevents any thorough adjustment of results with various demographic factors,⁴² such as breastfeeding, age at solid food introduction, number of siblings, day care attendance, exposure to smoking and pets, living on a farm, land use and mode of delivery (Table S3). Especially, factors associated with the living environment and lifestyle are very likely to affect individual skin/nasal microbiota.^{43,44} While we cannot establish causalities in this study, we feel that previous experimental work^{7,26,27} provides a solid mechanistic foundation that supports our results.

To conclude, in this nested case-control study from DIABIMMUNE infants with signs of allergic sensitization at the age of 18 months, we described differing nasal and skin microbiotas between Finland and Estonia. Namely, *Acinetobacter lwoffii* was more abundant and diverse in Estonian infants already at the age of 6 months, which could provide one potential explanation for the allergy gap between these populations. Differences in gut microbiota or between children later developing allergen-specific IgE responses were less clear. Possibly, studies from different time points of sample collection or a longer follow-up to detect development of the allergic phenotype could provide a more conclusive answer, regarding the role of the early commensal microbiota in the development of allergic sensitization.

ACKNOWLEDGMENTS

The authors thank the participating children and families in the DIABIMMUNE study as well as the physicians, nurses and technicians in the DIABIMMUNE study centres and laboratories. This work was supported by the European Union Seventh Framework Program (grant no. 202063), the Academy of Finland Centre of Excellence in Molecular Systems Immunology and Physiology Research (grant no. 250114) and the Liv och Hälsa Fund.

CONFLICTS OF INTEREST

The authors declare no conflict of interests.

AUTHOR CONTRIBUTIONS

LR, LvH, TH and MK conceived the study. LR, AP, AK and HS analysed the data. APe, AMH, VT, KK, SMV and ON contributed to data collection. LR and AP wrote the manuscript. All authors were involved in editing and final approval of the paper.

ORCID

Lasse Ruokolainen  <https://orcid.org/0000-0003-0951-9100>

Tari Haahtela  <https://orcid.org/0000-0003-4757-2156>

Mikael Knip  <https://orcid.org/0000-0003-0474-0033>

REFERENCES

1. Haahtela T. A biodiversity hypothesis. *Allergy*. 2019;74:1445-1456.
2. Ball TM, Castro-Rodriguez JA, Griffith KA, Holberg CJ, Martinez FD, Wright AL. Siblings, day-care attendance, and the risk of asthma and wheezing during childhood. *N Engl J Med*. 2000;343:538-543.
3. Riedler J, Braun-Fahrlander C, Eder W, et al. Exposure to farming in early life and development of asthma and allergy: a cross-sectional survey. *Lancet*. 2001;358:1129-1133.
4. von Hertzen L, Mäkelä MJ, Petäys T, et al. Growing disparities in atopy between the Finns and the Russians: A comparison of 2 generations. *J Allergy Clin Immunol*. 2006;117:151-157.
5. Ruokolainen L, Fyhrquist N, Haahtela T. The rich and the poor: environmental biodiversity protecting from allergy. *Curr Opin Allergy Clin Immunol*. 2016;16:421-426.
6. Fujimura KE, Demoor T, Rauch M, et al. House dust exposure mediates gut microbiome Lactobacillus enrichment and airway immune defense against allergens and virus infection. *Proc Natl Acad Sci*. 2014;111:805-810.
7. Debarry J, Garn H, Hanuszkiewicz A, et al. *Acinetobacter lwoffii* and *Lactococcus lactis* strains isolated from farm cowsheds possess strong allergy-protective properties. *J Allergy Clin Immunol*. 2007;119:1514-1521.
8. Ruokolainen L, Paalanen L, Karkman A, et al. Significant disparities in allergy prevalence and microbiota between the young people in Finnish and Russian Karelia. *Clin Exp Allergy*. 2017;47:665-674.
9. Kondrashova A, Mustalahti K, Kaukinen K, et al. Lower economic status and inferior hygienic environment may protect against celiac disease. *Ann Med*. 2008;40:223-231.
10. Kondrashova A, Reunanen A, Romanov A, et al. A six-fold gradient in the incidence of type 1 diabetes at the eastern border of Finland. *Ann Med*. 2005;37:67-72.
11. The DIAMOND Project Group. Incidence and trends of childhood type 1 diabetes worldwide 1990-1999. *Diabet Med*. 2006;23:857-866.
12. Simre K, Uibo O, Peet A, et al. Exploring the risk factors for differences in the cumulative incidence of coeliac disease in two neighboring countries: the prospective DIABIMMUNE study. *Dig Liver Dis*. 2016;48:1296-1301.
13. Björkstén B, Dumitrascu D, Foucard T, et al. Prevalence of childhood asthma, rhinitis and eczema in Scandinavia and Eastern Europe. *Eur Respir J*. 1998;12:432.
14. Patel SP, Järvelin M-R, Little MP. Systematic review of worldwide variations of the prevalence of wheezing symptoms in children. *Environ Health*. 2008;7:57.
15. Voor T, Julge K, Böttcher MF, Jenmalm MC, Duchén K, Björkstén B. Atopic sensitization and atopic dermatitis in Estonian and Swedish infants. *Clin Exp Allergy*. 2005;35:153-159.

16. Vatanen T, Kostic AD, d'Hennezel E, et al. Variation in microbiome LPS immunogenicity contributes to autoimmunity in humans. *Cell*. 2016;165:842-853.
17. Karkman A, Lehtimäki J, Ruokolainen L. The ecology of human microbiota: dynamics and diversity in health and disease. *Ann NY Acad Sci*. 2017;1399:78-92.
18. Arrieta M-C, Stiemsma LT, Dimitriu PA, et al. Infancy microbial and metabolic alterations affect risk of childhood asthma. *Sci Transl Med*. 2015;7:307ra152.
19. Ruokolainen L, Hertzen L, Fyhrquist N, et al. Green areas around homes reduce atopic sensitization in children. *Allergy*. 2015;70:195-202.
20. Liaw A, Wiener M. Classification and regression by randomForest. *R News*. 2002;2:18-22.
21. Kuhn M. Caret: classification and regression training;2017. <https://CRAN.R-project.org/package=caret>
22. Seifert H, Dijkshoorn L, Gerner-Smidt P, Pelzer N, Tjernberg I, Vaneechoutte M. Distribution of *Acinetobacter* species on human skin: comparison of phenotypic and genotypic identification methods. *J Clin Microbiol*. 1997;35:2819-2825.
23. Berlau J, Aucken H, Malnick H, Pitt T. Distribution of *Acinetobacter* species on skin of healthy humans. *Eur J Clin Microbiol Infect Dis*. 1999;18:179-183.
24. Vanbroekhoven K, Ryngaert A, Wattiau P, De Mot R, Springael D. *Acinetobacter* diversity in environmental samples assessed by 16S rRNA gene PCR-DGGE fingerprinting. *FEMS Microbiol Ecol*. 2004;50:37-50.
25. Ege MJ, Mayer M, Schwaiger K, et al. Environmental bacteria and childhood asthma. *Allergy*. 2012;67:1565-1571.
26. Fyhrquist N, Ruokolainen L, Suomalainen A, et al. *Acinetobacter* species in the skin microbiota protect against allergic sensitization and inflammation. *J Allergy Clin Immunol*. 2014;134:1301-1309.
27. Conrad ML, Ferstl R, Teich R, et al. Maternal TLR signaling is required for prenatal asthma protection by the nonpathogenic microbe *Acinetobacter lwoffii* F78. *J Exp Med*. 2009;206:2869-2877.
28. Brand S, Teich R, Dicke T, et al. Epigenetic regulation in murine offspring as a novel mechanism for transmaternal asthma protection induced by microbes. *J Allergy Clin Immunol*. 2011;128:618-625.
29. Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature*. 2012;486:207-214.
30. Abrahamsson TR, Jakobsson HE, Andersson AF, Björkstén B, Engstrand L, Jenmalm MC. Low gut microbiota diversity in early infancy precedes asthma at school age. *Clin Exp Allergy*. 2014;44:842-850.
31. Laursen MF, Bahl MI, Michaelsen KF, First LTR. Foods and Gut Microbes. *Front Microbiol*. 2017;8:356. <https://doi.org/10.3389/fmicb.2017.00356>.
32. Ranucci G, Buccigrossi V, Borgia E, et al. Galacto-Oligosaccharide/Polidextrose enriched formula protects against respiratory infections in infants at high risk of atopy: a randomized clinical trial. *Nutrients*. 2018;10:286.
33. Trompette A, Gollwitzer ES, Yadava K, et al. Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. *Nat Med*. 2014;20:159-166.
34. Lynch SV, Wood RA, Boushey H, et al. Effects of early-life exposure to allergens and bacteria on recurrent wheeze and atopy in urban children. *J Allergy Clin Immunol*. 2014;134(593-601):e12.
35. Brockow I, Zutavern A, Hoffmann U, et al. Early allergic sensitizations and their relevance to atopic diseases in children aged 6 years: results of the GINI study. *J Investig Allergol Clin Immunol*. 2009;19:180-187.
36. von Hertzen L, Haahtela T. Con: house dust mites in atopic diseases. *Am J Respir Crit Care Med*. 2009;180:113-119.
37. Silvers KM, Frampton CM, Wickens K, et al. Breastfeeding Protects against Current Asthma up to 6 Years of Age. *J Pediatr*. 2012;160:991-996.
38. Ek WE, Karlsson T, Hernández CA, Rask-Andersen M, Johansson Å. Breast-feeding and risk of asthma, hay fever, and eczema. *J Allergy Clin Immunol*. 2018;141:1157-1159.
39. Garcia-Larsen V, Ierodiakonou D, Jarrold K, et al. Diet during pregnancy and infancy and risk of allergic or autoimmune disease: A systematic review and meta-analysis. *PLoS Med*. 2018;15:e1002507.
40. Bisgaard H, Li N, Bonnelykke K, et al. Reduced diversity of the intestinal microbiota during infancy is associated with increased risk of allergic disease at school age. *J Allergy Clin Immunol*. 2011;128:646-652.
41. Kallionpää H, Laajala E, Öling V, et al. Standard of hygiene and immune adaptation in newborn infants. *Clin Immunol*. 2014;155:136-147.
42. Hanski I, von Hertzen L, Fyhrquist N, et al. Environmental biodiversity, human microbiota, and allergy are interrelated. *Proc Natl Acad Sci*. 2012;109:8334-8339.
43. Hakanen E, Lehtimäki J, Salmela E, et al. Urban environment predisposes dogs and their owners to allergic symptoms. *Sci Rep*. 2018;8:1585.
44. Lehtimäki J, Sinkko H, Hielm-Björkman A, et al. Skin microbiota and allergic symptoms associate with exposure to environmental microbes. *Proc Natl Acad Sci*. 2018;115:4897-4902.

SUPPORTING INFORMATION

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

How to cite this article: Ruokolainen L, Parkkola A, Karkman A, et al. Contrasting microbiotas between Finnish and Estonian infants: Exposure to *Acinetobacter* may contribute to the allergy gap. *Allergy*. 2020;00:1-10. <https://doi.org/10.1111/all.14250>

Supporting information

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to:

Lasse Ruokolainen. Anna Parkkola. Antti Karkman. Hanna Sinkko. Aleksandr Peet. Anu-Maaria Hämäläinen.

Leena von Herzen. Vallo Tillmann. Katriina Koski. Suvi M. Virtanen. Onni Niemelä. Tari Haahtela. Mikael

Knip. **Contrasting microbiotas between Finnish and Estonian infants: exposure to *Acinetobacter* may explain the allergy gap.** *Allergy*.

-Supplemental Material and methods

-Supplemental Figure S1

-Table S1

-Table S2

-Table S3

Supplemental material and methods:

1 Sample collection

The DIABIMMUNE birth cohort prospectively followed children with HLA susceptibility to type 1 diabetes. Newborn infants of consenting parents were screened in three centers: Espoo (Finland), Tartu (Estonia), and Petrozavodsk (Russia). HLA eligibility was defined as having high [DRB1*03-DQA1*05-DQB1*02 (DR3-DQ2)/DRB1*0401/2/4/5-DQA1*03-DQB1*03:02 (DR4-DQ8) genotype], moderate (DR4-DQ8/X genotype; X = non-protective allele but not DR3-DQ2) or slightly increased (DR3-DQ2/Y genotype; Y = non-protective allele but not DR4-DQ8) HLA class II conferred risk for type 1 diabetes¹. Among the screened newborn infants in Espoo, 23% were eligible by HLA, whereas the proportion was 16% in Tartu².

At study visits (3, 6, 12, 18, 24, and 36 months of age), nasal and skin swabs were collected, whereas stool samples were collected at home once a month. Skin swabs were taken from the volar surface of the forearm midway between wrist and elbow with soaked sterile cotton swabs. Nasal samples were taken by rotating a cotton swab for five seconds in the left nostril. All samples were initially frozen at -20°C and transported to be stored at -70°C. At 18 months of age, total IgE and allergen-specific IgE responses to seven allergens were tested by the ImmunoCAP fluoroenzyme immunoassay. The child's life events were recorded continuously by the parents in the DIABIMMUNE Study Book, including information on the infant's diet and exposure to animals (indoor/outdoor cat, dog, or other animal, living on a farm, visiting barns, contact with horse, cow, pig, sheep, chicken, other animals). Exclusive breastfeeding was defined as the infant receiving only breast milk, not ever having received any kind of formulas.

2 DNA extraction and sequencing

DNA was extracted from skin and nasal swabs using FastDNA™ SPIN Kit for Soil (MP Biomedicals, Santa Ana, CA) and from faecal samples with QIAamp DNA Stool Mini Kit (Qiagen, Hilden, Germany). DNA was used as a template to amplify V1-V3 regions of bacterial 16S rRNA genes by polymerase chain reaction (PCR) using 27F (AGAGTTTGATCMTGGCTCAG)³ and pD' (GTATTACCGCGGCTGCTG)⁴ primers with adapter sequences. Thermal cycles in PCR were performed as previously described by Lehtimäki et al.⁵ Three

parallel PCR products were pooled and sequenced with Illumina® Miseq runs at the Institute of Molecular Medicine Finland (FIMM), University of Helsinki.

3 IgE measurements

Concentrations of specific IgE antibodies were measured in a SFS-EN ISO/IEC 17025:2005- and SFS-EN ISO 15189:2007- accredited centralized clinical chemistry laboratory by using an automated Phadia 250 ImmunoCAP fluoroenzyme immunoassay analyzer (Phadia Diagnostics, Thermo Fisher Scientific). The analyses were carried out blindly to the knowledge of the clinical and demographic data.

4 Bioinformatics

PCR primers were removed using cutadapt v. 1.10⁶ and paired end sequencing reads merged with PEAR v. 0.9.6.⁷. Reads were quality trimmed using vsearch v. 2.0.2.⁸ with command fastq_filter and parameters fastq_maxee 3. fastq_minlen 365. fastq_maxns 0. Next, chimeric sequences were removed using the uchime_ref command, against the Silva-gold set v.123 (available from https://www.mothur.org/wiki/Silva_reference_files).

After quality trimming, reads were dereplicated and clustered with abundance based greedy clustering algorithm as implemented in vsearch v.2.0.2. Representative sequences from the resulting OTUs were aligned and taxonomically annotated with mothur v. 1.38.1.1⁹ against the Silva v.123 database (<https://www.arb-silva.de/documentation/release-123>). Phylogenetic tree was drawn using FastTree v. 2.1.8.¹⁰ using the GTR model.

Contaminant OTUs were identified by comparing samples to controls (both negative controls, i.e., blank samples, and PCR controls were used) and manually removed. Also, samples with library sizes smaller than 4,000 reads were removed. Due to sequencing run batch effects, OTUs that had less than 10 reads and not shared with both sequencing runs were removed. After contaminant removal, the sequencing libraries were normalized using the CSS method with log2 transformation as implemented in metagenomeSeq

package v.1.11.16¹¹ in R v.3.1.1¹². Oligotyping of OTUs was done using all reads assigned to an OTU by vsearch clustering using oligotyping pipeline v.2.1.¹³

To analyse community composition, we calculated the Bray&Curtis dissimilarity¹⁴ and weighted UniFrac distance¹⁵. Between-sample community dissimilarity was compared using permutational multiple regression on distance matrices (MRM), as implemented in the vegan package in R¹⁶. Statistical significance was estimated using 999 random permutations. Between-group comparisons on diversity or OTU abundances was done with generalized least squares, assuming groups-specific residual dispersion, as implemented in the nlme package¹⁷.

References

1. Peet A, Hämäläinen A-M, Kool P, Ilonen J, Knip M, Tillmann V et al. Early postnatal growth in children with HLA-conferred susceptibility to type 1 diabetes. *Diabetes Metab Res Rev* 2014;30:60–68.
2. Mustonen N, Siljander H, Peet A, Tillmann V, Harkonen T, Ilonen J et al. Early childhood infections precede development of beta-cell autoimmunity and type 1 diabetes in children with HLA-conferred disease risk. *Pediatr Diabetes* 2018;19:293–299.
3. Chung J, Ha E-S, Park H-R, Kim S. Isolation and characterization of Lactobacillus species inhibiting the formation of Streptococcus mutans biofilm. *Mol Oral Microbiol* 2004;19:214–216.
4. Edwards U, Rogall T, Blöcker H, Emde M, Böttger EC. Isolation and direct complete nucleotide determination of entire genes. Characterization of a gene coding for 16S ribosomal RNA. *Nucleic Acids Res* 1989;17:7843–7853.
5. Lehtimäki J, Karkman A, Laatikainen T, Paalanen L, Hertzen L von, Haahtela T et al. Patterns in the skin microbiota differ in children and teenagers between rural and urban environments. *Sci Rep* 2017;7:45651.
6. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 2011;17:pp–10.
7. Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 2013;30:614–620.
8. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 2016;4:e2584.
9. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 2009;75:7537–7541.
10. Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 2009;26:1641–1650.

11. Paulson JN. Stine OC. Bravo HC. Pop M. Differential abundance analysis for microbial marker-gene surveys. *Nat Methods* 2013;10:1200–1202.
12. R Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing 2016
13. Eren AM. Maignien L. Sul WJ. Murphy LG. Grim SL. Morrison HG et al. Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods Ecol Evol* 2013;4:1111–1119.
14. Legendre P. Legendre LF. *Numerical ecology*. Elsevier 2012
15. Lozupone C. Lladser ME. Knights D. Stombaugh J. Knight R. UniFrac: an effective distance metric for microbial community comparison. *ISME J Lond* 2011;5:169–172.
16. Oksanen J. Blanchet G. Friedly M. Kindt R. Legendre P. McGlinn D et al. vegan: Community ecology package. R package version 2.4-3. 2017.<https://CRAN.R-project.org/package=vegan>
17. Pinheiro J. Bates D. DebRoy S. Sarkar D. R Core Team. nlme: Linear and nonlinear mixed effects Models. R package version 3.1-131. 2017.<https://CRAN.R-project.org/package=nlme>

Supplemental Figure S1

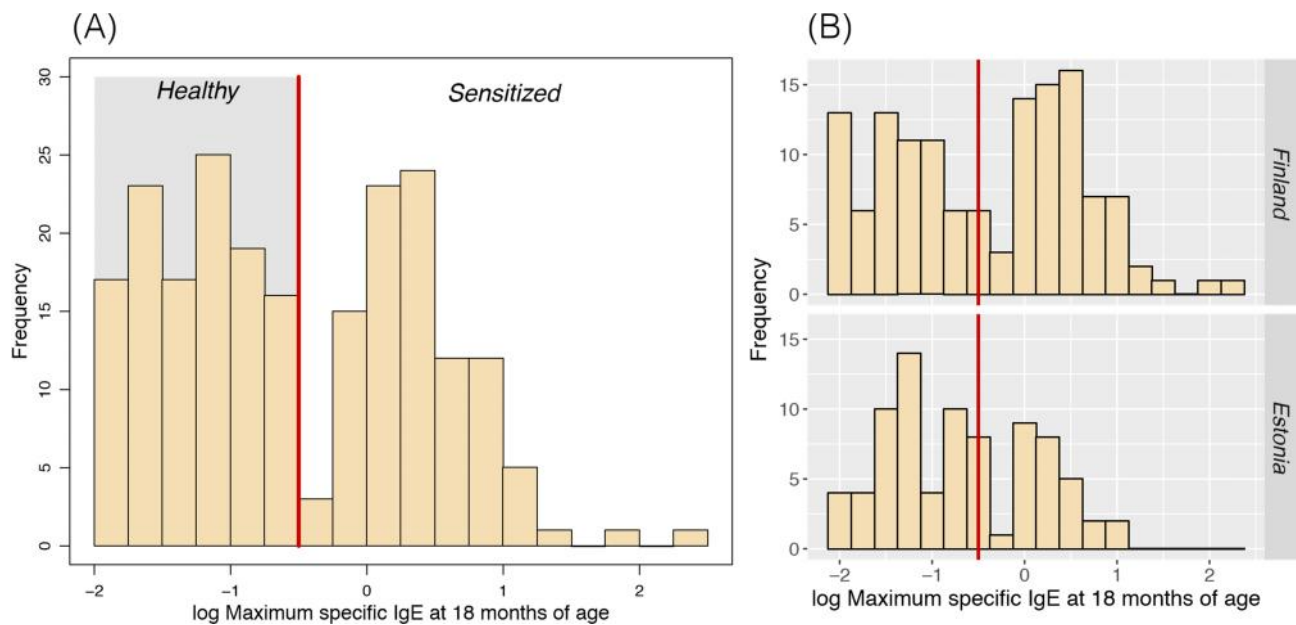


Figure S1. Classification of subjects was done based on specific IgE measurements at the age of 18 months. (A) IgE cut-off was set empirically at -0.5 on log-scale (red vertical lines). based on the distribution of the maximum specific IgE value across populations. (B) Population-specific distributions of the maximum specific IgE value. along with the chosen cut-off value.

Table S1: Proportion of Finnish and Estonian study participants with positive sIgE responses (≥ 0.35 kU/L) to the seven allergens tested at the age of 18 months.

Allergen	Finnish children (n=140)	Estonian children (n=84)
Cat, n (%)	9 (6.4)	7 (8.3)
Hen's egg, n (%)	52 (37.1)	17 (20.5)
Cow's milk, n (%)	44 (31.4)	18 (22.5)
Peanut, n (%)	11 (8.1)	3 (4.3)
Birch, n (%)	9 (6.4)	0
Timothy, n (%)	3 (2.2)	0
Dust mite, n (%)	2 (1.4)	1 (1.3)

Table S2: Demographic and other characteristics of the Finnish and Estonian study participants.

Characteristic	Finland (n=140)	Estonia (n=84)
Boys, n (%)	76 (54.3)	41 (48.8)
Number of siblings, median (range)	1 (0-5)	1 (0-6)
Caesarean section, n (%)	11 (7.9)	9 (10.7)
Duration of exclusive breast-feeding, months, mean (range)	3.5 (0-7.0)	1.6 (0-6.4)
Duration of total breastfeeding, months, mean (range)	9.9 (0.2-28.0)	9.6 (0-29.8)
Age at introduction of solid food, months, mean (range)	4.4 (2.5-7.0)	4.2 (0-13)
Maternal smoking, n (%)*	12 (8.6)	7 (8.3)
Paternal smoking, n (%)*	26 (18.8)	30 (35.7)
Day care attendance, n (%)*	87 (62.1)	64 (76.2)
Age at start of day care, months, mean (range)	19.4 (2-39)	25.5 (11-37)
Indoor pets in the family, n(%)*	48 (34.3)	46 (54.8)
Living on a farm, n (%)*	0 (0)	15 (17.9)
Proportion of green environment (forest+agricultural), mean (range)	0.43 (0.14–0.81)	0.61 (0.14–1)

*At any timepoint up to the 3-year study visit

Table S3: Demographic and other characteristics of sensitised and non-sensitised participants.

Characteristic	Sensitised (n=98)	Non-sensitised (n=126)
Boys, n (%)	58 (59.2)	59 (46.8)
Number of siblings, median (range)	1 (0-5)	1 (0-6)
Caesarean section, n (%)	6 (6.1)	14 (11.1)
Duration of exclusive breastfeeding, months, mean (range)	2.7 (0-7)	2.7 (0-6.4)
Duration of total breastfeeding, months, mean (range)	10.2 (0.1-29.8)	9.5 (0-28.7)
Age at introduction of solid food, months, mean (range)	4.4 (2-13)	4.2 (0-6.0)
Maternal smoking, n (%)*	4 (4.1)	15 (11.9)
Paternal smoking, n (%)*	16 (16.3)	40 (31.7)
Day care attendance, n (%)*	67 (68.4)	84 (66.7)
Age at start of day care, months, mean (range)	20.7 (9-37)	23.0 (2-39)
Indoor pets in the family, n(%)*	43 (43.9)	51 (40.5)
Living on a farm, n (%)*	4 (4.1)	11 (8.7)
Land use gradient from forested to built environment, mean (range)	0.49 (0.14-1)	0.51 (0.14-1)

*At any timepoint up to the 3-year study visit